



# Trans-ancestry meta-analyses identify rare and common variants associated with blood pressure and hypertension

## Citation

Surendran, P., F. Drenos, R. Young, H. Warren, J. P. Cook, A. K. Manning, N. Grarup, et al. 2016. "Trans-ancestry meta-analyses identify rare and common variants associated with blood pressure and hypertension." *Nature genetics* 48 (10): 1151-1161. doi:10.1038/ng.3654. <http://dx.doi.org/10.1038/ng.3654>.

## Published Version

doi:10.1038/ng.3654

## Permanent link

<http://nrs.harvard.edu/urn-3:HUL.InstRepos:32072248>

## Terms of Use

This article was downloaded from Harvard University's DASH repository, and is made available under the terms and conditions applicable to Other Posted Material, as set forth at <http://nrs.harvard.edu/urn-3:HUL.InstRepos:dash.current.terms-of-use#LAA>

## Share Your Story

The Harvard community has made this article openly available.  
Please share how this access benefits you. [Submit a story](#).

[Accessibility](#)

Published in final edited form as:

*Nat Genet.* 2016 October ; 48(10): 1151–1161. doi:10.1038/ng.3654.

## Trans-ancestry meta-analyses identify rare and common variants associated with blood pressure and hypertension

Users may view, print, copy, and download text and data-mine the content in such documents, for the purposes of academic research, subject always to the full Conditions of use: [http://www.nature.com/authors/editorial\\_policies/license.html#terms](http://www.nature.com/authors/editorial_policies/license.html#terms)

Corresponding authors: Joanna M M Howson, [jmmh2@medschl.cam.ac.uk](mailto:jmmh2@medschl.cam.ac.uk) and Patricia B Munroe, [p.b.munroe@qmul.ac.uk](mailto:p.b.munroe@qmul.ac.uk).

<sup>168</sup>Jointly supervised the work

### URLs

Exome chip design information: [http://genome.sph.umich.edu/wiki/Exome\\_Chip\\_Design](http://genome.sph.umich.edu/wiki/Exome_Chip_Design)

RareMetalWorker information: <http://genome.sph.umich.edu/wiki/RAREMETALWORKER>

Summary SNV association results: <http://www.phenoscaner.medschl.cam.ac.uk>

Databases used for variant annotation: <http://www.ncbi.nlm.nih.gov/SNP/> <http://www.ensembl.org/info/docs/tools/index.html> and <http://evs.gs.washington.edu/EVS/>

UCSC reference file used for annotation of variants with gene and exon information: <http://hgdownload.soe.ucsc.edu/goldenPath/hg19/database/refFlat.txt.gz>

Databases used for pathway analysis: MAGENTA (<https://www.broadinstitute.org/mpg/magenta/>) and THOMSON REUTERS MetaCore™ Single Experiment Analysis workflow tool (<http://thomsonreuters.com/en/products-services/pharma-life-sciences/pharmaceutical-research/metacore.html>).

### Author contributions

Supervision and management of the project: JMMH and PBM. The following authors contributed to the drafting of the manuscript: JMMH, PBM, PSu, HW, ASB, FD, JPC, DRB, KW, MT, FWA, LVW, NJS, JD AKM, HY, CMM, NG, XS, TaT, DFF, MHs, OG, TF, VT. All authors critically reviewed and approved the final version of the manuscript. **Statistical analysis review:** JMMH, PSu, FD, HW, JPC, RY, NM, PBM, LVW, HY, TF, EMi, ADM, AM, AM, EE, ASB, FWA, MJC, CF, TF, SEH, ASH, JEH, JL, GM, JM, NM, APM, APo, NJS, RAS, LS, KE, MT, VT, TVV, NV, KW, AMY, WZg, NG, CML, AKM, XS, TT. **Central Data QC:** JMMH, ASB, PSu, RY, FD, HW, JPC, TF, LVW, PBM, EMi, NM, CML, NG, XS, AKM. **Central Data analysis:** JMMH, PSu, FD, HW, JPC, NG, CML, AKM, XS. **Pathway analysis and literature review:** JMMH, DRB, PBM, MT, KW, VT, OG, AT, FWA. **GWAS lookups, eQTL analysis, GRS, variant annotation and enrichment analyses:** JMMH, ASB, DRB, JRS, DFF, FD, MHR, PBM, FWA, TT, CML, AKM, SBU. **Study Investigators in alphabetical order by consortium (CHD Exome+, ExomeBP and GoT2D):** DSA, PA, EA, DA, ASB, RC, JD, JF, IF, PF, JWJ, FKe, ASM, SFN, BGN, DS, NSa, JV, FWA, PIWB, MJB, MJC, JCC, JMC, IJD, GD, AFD, PE, TE, PWF, GG, PH, CH, KH, EI, MJ, FKa, SK, JSK, LLi, MIM, OM, AMe, ADM, APM, PBM, MEN, SP, CP, OPo, DP, SR, OR, IR, VS, NJS, PSe, TDS, JMS, NJW, CJW, EZ, MB, IB, FSC, LG, TH, EKH, PJ, JKu, ML, TAL, AL, KLM, HO, OPe, RR, JT, MU. **Study Phenotyping in alphabetical order by consortium (CHD Exome+, ExomeBP and GoT2D):** PA, DA, SBI, MC, JF, JWJ, FKe, KK, SFN, BGN, CJP, AR, MS, NSa, JV, WZo, RAB, MJB, MJC, JCC, JMC, AFD, ASFD, LAD, TE, AF, GG, GH, PH, AS H, OLH, EI, MJ, FK, JSK, LLi, LLa, GM, AMc, PM, AMe, RMg, MJN, MEN, OPo, NP, FR, VS, NJS, TDS, AVS, JMS, MT, AV, NV, NJW, TIT, CC, LLH, MEJ, AK, PK, JL DPS, SM, ERBP, AS, TS, HMS, BT. **Study Data QC and analysis in alphabetical order by consortium (CHD Exome+, ExomeBP and GoT2D):** ASB, AJMC, JMMH, JK, SFN, BGN, MMN, SP, MP, PSu, ST, GV, SMW, RY, FWA, JPC, FD, AF, TF, CH, AMc, AMj, APM, PBM, CP, WR, FR, NJS, MT, VT, HW, HY, NG, AKM, XS. **Exome chip data QC in alphabetical order by consortium (CHD Exome+, ExomeBP and GoT2D):** ASB, JMMH, SFN, BGN, PSu, RY, FWA, PIWB, AIFB, JCC, JPC, PD, LAD, FD, EE, CF, TF, SEH, PH, SSH, KH, JEH, EK, AMj, GM, JM, NM, EMi, AMo, APM, PBM, CPN, MJN, CP, AP, WR, NRR, RAS, NS, LS, KES, MDT, VT, TVV, TVV, NV, HW, HY, AMY, EZ, WZg, NG, CML, AKM, XS. **Exome chip Data analysis in alphabetical order by consortium (CHD Exome+, ExomeBP and GoT2D):** JMMH, PSu, RY, FWA, PIWB, AIFB, RAB, MJC, JCC, JPC, PD, LAD, PE, EE, CF, TF, PWF, SF, CG, SEH, PH, ASH, CH, OLH, JEH, EI, MJ, FKa, JSK, DCML, LLi, JL, GM, RMR, JM, NM, MIM, PM, OM, CM, EMi, AMo, APM, RMg, PBM, CPN, MJN, TO, APo, APa, WR, NRR, NJS, RAS, NS, LS, TDS, KES, MDT, ET, VT, TVV, NV, LVW, NJW, HW, HY, AMY, EZ, HZ, WZg, LLB, APG, NG, MHs, JRH, AUJ, JBJ, CML, AKM, NN, XS, AS, AJS. **GRS lookups:** AEJ, EMa, HFM, HL, HMH, JFF, MTr, RSV, WL.

### Conflict of interests

N. P. has received financial support from several pharmaceutical companies that manufacture either blood pressure lowering or lipid lowering agents, or both, and consultancy fees.

S. K. has received Research Grant-Merck, Bayer, Aegerion; SAB-Catabasis, Regeneron Genetics Center, Merck, Celera; Equity-San Therapeutics, Catabasis; Consulting-Novartis, Aegerion, Bristol Myers-Squibb, Sanofi, AstraZeneca, Alnylam.

P. Sever has received research awards from Pfizer Inc.

A. Malarstig and M. Uria-Nickelsen are full time employees of Pfizer.

D. Reily and M. Hoek are full time employees of Merck and co Inc.

M.J. Caulfield is Chief Scientist for Genomics England a UK Government company.

The authors declare no competing financial interest.

*A full list of authors and affiliations appears at the end of the article.*

# These authors contributed equally to this work.

## Abstract

High blood pressure is a major risk factor for cardiovascular disease and premature death. However, there is limited knowledge on specific causal genes and pathways. To better understand the genetics of blood pressure, we genotyped 242,296 rare, low-frequency and common genetic variants in up to ~192,000 individuals, and used ~155,063 samples for independent replication. We identified 31 novel blood pressure or hypertension associated genetic regions in the general population, including three rare missense variants in *RBM47*, *COL21A1* and *RRAS* with larger effects ( $>1.5\text{mmHg/allele}$ ) than common variants. Multiple rare, nonsense and missense variant associations were found in *A2ML1* and a low-frequency nonsense variant in *ENPEP* was identified. Our data extend the spectrum of allelic variation underlying blood pressure traits and hypertension, provide new insights into the pathophysiology of hypertension and indicate new targets for clinical intervention.

## Introduction

High blood pressure (BP) or hypertension is a highly prevalent chronic disorder. It is estimated to be responsible for a larger proportion of global disease burden and premature mortality than any other disease risk factor<sup>1</sup>. Elevated systolic and/or diastolic BP increases the risk of several cardiovascular disorders including stroke, coronary heart disease (CHD), heart failure, peripheral arterial disease and abdominal aortic aneurysms<sup>2</sup>. BP is a complex, heritable, polygenic phenotype for which genome-wide association studies (GWAS) have identified over 67 genetic regions associated with BP and/or hypertension to date<sup>3–11</sup>. These variants are common (minor allele frequency, MAF  $>0.05$ ), mostly map to intronic or intergenic regions, with the causal alleles and genes not readily identified due to linkage disequilibrium (LD)<sup>4,5</sup>, and explain only ~2% of trait variance<sup>12</sup>. Low-frequency ( $0.01 < \text{MAF} < 0.05$ ) and rare (MAF  $< 0.01$ ) single nucleotide variants (SNVs), predominantly unexplored by GWAS may have larger phenotypic effects than common SNVs<sup>13</sup>, and may help to explain the missing heritability, and identify causative genes as demonstrated previously<sup>14</sup>.

To identify novel coding variants and loci influencing BP traits and hypertension we performed the largest meta-analysis to date that included a total of ~350,000 individuals, directly genotyped with the Exome chip. The Exome chip contains ~240,000 mostly rare and low-frequency variants (Methods). A single-variant discovery analysis was performed, and candidate SNVs were taken forward for validation using independent replication samples. Gene-based tests were used to identify BP associated genes harboring multiple rare variant associations. We next assessed whether the newly identified BP associated SNVs were associated with expression levels of nearby genes, and tested these variants in aggregate for a causal association of BP with other cardiovascular traits and risk factors. Our findings highlight the contribution of rare variants in the aetiology of blood pressure in the general population, and provide new insights into the pathophysiology of hypertension.

## Results

### Discovery of single variant BP associations

We genotyped 192,763 individuals from 51 studies, and assessed association of 242,296 SNVs with diastolic BP (DBP), systolic BP (SBP), pulse pressure (PP) and hypertension (HTN; Supplementary Tables 1, 2 and 3; Methods). An overview of the SNV discovery study design is given in Figure 1. A fixed effects meta-analysis for each trait was performed using study-level association summary statistics from i) samples of European (EUR) ancestry (up to 165,276 individuals), and ii) a trans-ethnic meta-analysis of the EUR and additional South Asian (SAS) ancestry samples (EUR\_SAS; up to 192,763 individuals). Two analyses of DBP, SBP and PP were performed, one in which the trait was inverse normal transformed and a second in which the raw phenotype was analysed. Both sets of results were consistent (Methods), therefore to minimise sensitivity to deviations from normality in the analysis of rare variants, the results from the analyses of the transformed traits were used for discovery. Strong correlations between the BP traits were observed across studies (Methods), hence no adjustment of significance thresholds for independent trait testing was applied.

The discovery meta-analyses identified 51 genomic regions with genome-wide significant (GWS) evidence of association with at least one of the four BP traits tested ( $P < 5 \times 10^{-8}$ ; Supplementary Table 4). There were 46 regions associated in the EUR\_SAS samples, of which 14 were novel (Supplementary Figure 1). An additional five regions were GWS in the EUR only meta-analyses of which three were novel (Supplementary Figure 2). In total, 17 genomic regions were identified that were GWS for at least one BP trait that have not been previously reported.

### Replication of single variant BP associations

Next we sought support for our findings, in an independent replication dataset comprising of 18 studies, 15 of which were from the Cohorts for Heart and Aging Research in Genomic Epidemiology+ (CHARGE+) exome chip blood pressure consortium (Figure 1; Liu *et al.* Nature Genetics, *submitted*). Variants were selected for replication first using the larger (transformed) EUR\_SAS data, with additional variants from the (transformed) EUR data also selected. SNVs were selected if they mapped outside of known BP genomic regions and had  $MAF \geq 0.05$  and  $P < 1 \times 10^{-5}$  or  $MAF < 0.05$  and  $P < 1 \times 10^{-4}$  with at least one BP trait, *i.e.* choosing a lower significance threshold for the selection of rare variants (full details of the selection criteria are provided in the Methods). In total 81 candidate SNVs were selected for replication (Supplementary Table 5). Eighty variants were selected from EUR\_SAS (transformed) results and one SNV at the *ZNF101* locus from the EUR (transformed) analyses. The results for EUR\_SAS and EUR were consistent (association statistics were correlated,  $\rho = 0.9$  across ancestries for each of the traits). Of the 81 variants, 30 SNVs were selected for association with DBP as the primary trait, 26 for SBP, 19 for PP and 6 for HTN, with the primary trait defined as the BP trait with the smallest association *P*-value in the EUR-SAS discovery analyses.

Meta-analyses were performed on results from analyses of untransformed DBP, SBP, PP and HTN (as only results of untransformed traits were available from CHARGE+) in (i) up to 125,713 individuals of EUR descent, and (ii) up to 155,063 individuals of multiple ethnicities (4,632 of Hispanic descent, 22,077 of African American descent, 2,641 SAS samples with the remainder EUR; Figure 1). Given that a large proportion of the ancestries in the trans-ethnic meta-analyses were not included in our discovery samples, we used the EUR meta-analyses as the main data set for replication, but we also report any additional associations identified within the larger trans-ethnic dataset.

Novel BP-SNV associations were identified based on two criteria (Figure 1; Methods). Firstly, replication of the primary BP trait-SNV association was sought at a Bonferroni adjusted  $P$ -value threshold in the replication data ( $P \leq 6.17 \times 10^{-4}$ , assuming  $\alpha=0.05$  for 81 SNVs tested and same direction of effect; Methods) without the need for GWS. Secondly, meta-analyses of discovery and replication results across all four (untransformed) BP traits were performed to assess the overall level of support across all samples for the 81 candidate SNVs; those BP-SNV associations that were GWS (with statistical support in the replication studies;  $P < 0.05$  and the same direction of effect) were also declared as novel.

Seventeen SNV-BP associations formally replicated with concordant direction of effect at a Bonferroni adjusted significance level for the primary trait. Fourteen were in the EUR meta-analyses, and amongst these was a rare non-synonymous (ns) SNV mapping to *COL21A1* (Table 1, Supplementary Table 6). Three associations were in the trans-ethnic meta-analyses, these included two rare nsSNVs in *RBM47* and *RRAS* (Table 1, Supplementary Table 7; Methods).

In addition to the 17 SNV-BP trait associations that formally replicated, we identified 13 further SNV-associations that were GWS in the combined (discovery and replication) meta-analyses. Ten of these were GWS in the combined EUR analyses, (Table 2; Supplementary Tables 6 and 8a), and three were GWS in the combined trans-ethnic meta-analyses (Table 2; Supplementary Tables 7 and 8b).

This gives a total of 30 novel SNV-BP associations (15 SNV-DBP, 9 SNV-SBP and 6 SNV-PP; Tables 1 and 2; Supplementary Figures 3 and 4). Five of the SNVs were GWS with more than one BP trait (Figure 2: Tables 1 and 2; Supplementary Table 8). Four loci (*CERS5*, *TBX2*, *RGL3* and *OBFC1*) had GWS associations with HTN in addition to GWS associations with DBP and SBP. The *PRKAG1* locus had GWS associations with both SBP and PP.

Conditional analyses were performed to identify secondary signals of association within the novel BP loci. The RAREMETALWORKER (RMW) package (Methods)<sup>15</sup> allows conditional analyses to be performed using summary level data. Hence, analyses of the transformed primary traits and HTN were re-run in RMW across the discovery studies (Figure 3). The results of the RMW single variant tests were consistent with the initial discovery analyses (Supplementary Information). Given the RMW analyses were based on our discovery samples, the larger EUR-SAS data was used as the main analysis to increase power, but we also report any additional associations with evidence in EUR.

We identified secondary independent signals of association in four loci, *PREX1*, *PRKAG1* and *RRP1B* within the EUR\_SAS analyses and *COL21A1* in the EUR analyses ( $P_{\text{conditional}} < 1 \times 10^{-4}$ , Bonferroni adjusted for ~500 variants within each region; Methods; Supplementary Tables 9 and 10). Three independent association signals were identified in the *MYH6* locus in the EUR\_SAS analyses (Supplementary Table 11).

### Gene-based BP associations

To improve statistical power to detect associations in genes harbouring rare variants, analytical methods that combine effects of variants across a gene into a single test have been devised and are implemented in the RMW package<sup>15</sup>. We applied the gene-based sequence kernel association test (SKAT)<sup>16</sup> and Burden tests<sup>17</sup> to the RMW dataset (MAF < 0.05 or MAF < 0.01; Figure 3; Methods). One previously unidentified BP gene (*A2ML1*) was associated with HTN ( $P = 7.73 \times 10^{-7}$ ) in the EUR\_SAS studies and also in EUR studies (Supplementary Table 12; Bonferroni-corrected threshold of significance  $P < 2.8 \times 10^{-6}$ , after adjusting for 17,996 genes tested, Methods). The gene showed residual association with the primary BP trait after conditioning on the most associated SNV in the gene ( $P_{\text{conditional}} = 5.00 \times 10^{-4}$ ; Supplementary Table 12), suggesting that the association is due to multiple rare variants in the gene. One nonsense variant (rs199651558, p.Arg893\*, MAF =  $3.5 \times 10^{-4}$ ) was observed, and there were multiple missense variants (Figure 4). *A2ML1* encodes alpha-2-macroglobulin-like 1 protein, and is a member of the alpha macroglobulin superfamily, which comprises protease inhibitors targeting a wide range of substrates. Mutations in this gene are associated with a disorder clinically related to Noonan syndrome, a developmental disorder which involves cardiac abnormalities<sup>18</sup>. We sought replication in the CHARGE+ studies for this gene, however there was no evidence of association with HTN ( $P = 0.45$ ). Given the very low frequencies of the variants involved, however, studies in which the variants are polymorphic will be required to replicate the association with HTN. The *DBH* gene was found to be associated with DBP using the SKAT test ( $P = 2.88 \times 10^{-6}$ ). However, this was not due to multiple rare variants as the association was driven by rs77273740 (Supplementary Table 5) and the SNV was not validated in the replication samples.

### Rare and common variant associations in established BP loci

Of the 67 established BP loci, 35 loci were on the Exome chip (N=43 SNVs or close proxies  $r^2 > 0.7$ ). All 43 SNVs had at least nominal evidence of association with BP in our discovery samples ( $P < 0.01$ ; Supplementary Table 13). We also assessed if any of the established BP loci contained coding variants that are associated with BP traits and in LD ( $r^2 > 0.2$ ) with the known BP variants on the Exome chip (Supplementary Table 13), using the 1000G phase 3 release for LD annotation. Focusing on SNVs that were GWS for any BP trait from our transformed discovery data for either ancestry, there were 25 coding variants, of which 6 were predicted to be damaging at loci labelled *CDC25A*, *SLC39A8*, *HFE*, *ULK4*, *ST7L-CAPZA1-MOV10* and *CYP11A1-ULK3*. Three of these are published variants at loci labelled *SLC39A8*, *HFE* and *ST7-CAPZA1-MOV10*. At *CYP11A1-ULK3*, the coding variant was in moderate LD with the reported variant, but was less significantly associated with DBP in our EUR\_SAS dataset ( $P = 2.24 \times 10^{-8}$  compared to  $P = 1.68 \times 10^{-15}$  for the published variant). At the *ULK4* locus the predicted damaging coding variant had similar association



as the published coding variant (predicted to be benign), and prior work has already indicated several associated nsSNVs in strong LD in *ULK4* 19. The nsSNV within the *CDC25A* locus (rs11718350 in *SPINK8*) had similar association with DBP as the intergenic published SNV in our EUR\_SAS dataset ( $P=2.00 \times 10^{-8}$  compared to  $P=2.27 \times 10^{-8}$  for the published variant). Overall at least 5 of the known loci are consistent with having a coding causal variant.

Gene-based SKAT tests of all genes that map within 1 Mb of a previously reported SNV association (Supplementary Table 14), indicated no genes with multiple rare or low-frequency variant associations. Single variant conditional analyses showed that rs33966350, a rare nonsense variant in *ENPEP* (MAF=0.01) was associated with SBP ( $P_{\text{conditional}}=1.61 \times 10^{-5}$ ) in the EUR\_SAS samples (Supplementary Tables 14 and 15; Methods) independently of the known SNV (rs6825911). *ENPEP* encodes aminopeptidase A (APA) an enzyme of the renin-angiotensin-aldosterone system (RAAS) that converts angiotensin II (AngII) to AngIII.

There were no other established loci with convincing low-frequency or rare SNV associations in the EUR\_SAS samples. However, *HOXC4*, had evidence of a second independent signal with a rare missense SNV in EUR samples (rs78731604; MAF=0.005,  $P_{\text{conditional}}=5.76 \times 10^{-5}$ ; Supplementary Table 15). The secondary signal in the *HOXC4* region, mapped to *CALCOCO1*, ~300kb from the known SNV. The gene association (MAF 0.01,  $P=2.37 \times 10^{-5}$ ) was below the required significance threshold and attributable to rs78731604, which is not predicted to have detrimental effects on protein structure. Therefore, replication of this association is required. Three loci (*ST7L-CAPZA1-MOV10*, *FIGN-GRB14*, and *TBX5-TBX3*) had evidence of a second independent signal in the region in EUR\_SAS samples with a common variant ( $P_{\text{conditional}} < 1 \times 10^{-4}$ ; Supplementary Table 15) that has not been previously reported.

Having identified 30 novel loci associated with BP traits, as well as additional new independent SNVs at four novel loci and five known loci, we calculated the percent of the trait variance explained (Methods). This was 2.08%/2.11%/1.15% for SBP/DBP/PP for the 43 previously reported BP-SNVs covered in our dataset, increasing to 3.38%/3.41%/2.08% respectively with the inclusion of the 30 lead SNVs from novel loci, plus new independent SNV-BP associations identified from novel and known loci.

### Effect of BP SNVs on cardiovascular traits & risk factors

Amongst our novel BP-SNV associations, some have previously been reported to be associated with other cardiovascular traits and risk factors (Supplementary Table 16); these include coronary heart disease (CHD: *PHACTR1*, *ABO*)<sup>20,21</sup>, QT interval (*RNF207*)<sup>22</sup>, heart rate (*MYH6*)<sup>23</sup>, and cholesterol levels (2q36.3, *ABO*, *ZNF101*)<sup>24</sup>.

To test the impact of BP variants on cardiovascular endpoints and risk factors we created three weighted genetic risk scores (GRS) according to SBP/DBP/PP based on the newly identified and previously published BP variants (up to N=125; Methods). The GRS models were used to test the causal effect of BP on the following traits: ischemic stroke (including the subtypes, cardiometabolic, large and small vessel <sup>25</sup>), CHD, heart failure,<sup>26</sup> left

ventricular mass<sup>27</sup>, left ventricular wall thickness<sup>27</sup>, high-density lipoprotein cholesterol (HDL-c), low-density lipoprotein (LDL-c), triglycerides, total cholesterol, body mass index (BMI), waist-hip ratio adjusted BMI, height and estimated glomerular filtration rate (eGFR) (Methods). As expected, BP was positively associated with increased CHD risk (OR [95% CI]=1.39[1.22-1.59] per 10mmHg increase in SBP,  $P=6.07\times 10^{-7}$ ; 1.62[1.28-2.05] per 10mmHg increase in DBP,  $P=5.99\times 10^{-5}$ ; 1.70[1.34-2.16] per 10mmHg increase in PP,  $P=1.20\times 10^{-5}$ ; Table 3), and increased risk of ischemic stroke (OR [95% CI]=1.93[1.47-2.55] per 10mmHg increase in DBP,  $P=2.81\times 10^{-6}$ ; 1.57[1.35-1.84] per 10mmHg increase in SBP,  $P=1.16\times 10^{-8}$ ; 2.12[1.58-2.84] per 10mmHg increase in PP,  $P=5.35\times 10^{-7}$ ). The positive association with ischemic stroke was primarily due to large vessel stroke (Table 3). DBP and SBP were also positively associated with left ventricular mass (9.57 [3.98-15.17] gram increase per 10mmHg increase in DBP,  $P=8.02\times 10^{-4}$  and 5.13 [1.77-8.48] gram increase per 10mmHg increase in SBP,  $P=0.0027$ ) and left ventricular wall thickness (0.10 [0.06-0.13] cm increase per 10mmHg increase in DBP,  $P=1.88\times 10^{-8}$  and 0.05 [0.03-0.07] cm increase per 10mmHg increase in SBP,  $P=5.52\times 10^{-6}$ , Table 3). There was no convincing evidence to support the BP associated variants having an effect on lipid levels ( $P>0.1$ ), BMI ( $P>0.005$ ), waist hip ratio adjusted BMI ( $P>0.1$ ), height ( $P>0.06$ ), eGFR ( $P>0.02$ ) or heart failure ( $P>0.04$ ). The causal associations with CHD, stroke, and left ventricular measures augment the results from a previous association analysis using 29 BP variants<sup>28</sup>. Our data strongly support the previous observations of no causal relationship between BP and eGFR. Lack of evidence of a BP effect with heart failure may only be due to lack of power, as the association was in the expected direction.

### Possible functional variants at BP loci and candidate genes

Twenty-six of our newly discovered BP associated SNVs had MAF 0.05 and therefore due to extensive LD with other SNVs not genotyped on the Exome array, identifying the causal genes requires additional information. If a SNV is associated with increased or decreased expression of a particular gene, *i.e.* it is an expression quantitative trait locus (eQTL) this suggests the gene on which the SNV acts could be in the causal pathway. To help identify potential candidate causal genes in the novel BP loci (Supplementary Table 9), information from publicly available eQTL databases was investigated (MuTHER for LCL, adipose and skin and GTEx for nine tissues including the heart and tibial artery; Methods).

The DBP increasing allele of the nsSNV, rs7302981-A, was associated with increased expression of *CERS5* in: LCLs ( $P_{\text{MuTHER}}=3.13\times 10^{-72}$ ) skin ( $P_{\text{MuTHER}}=2.40\times 10^{-58}$ ) adipose ( $P_{\text{MuTHER}}=2.87\times 10^{-54}$ ) and nerve ( $P_{\text{GTEx}}=4.5\times 10^{-12}$ ) (Supplementary Figure 5). Additional testing (Methods) provided no evidence against colocalisation of the eQTL and DBP association signals, implicating *CERS5* as a candidate causal gene for this DBP locus. *CERS5* (LAG1 homolog, ceramide synthase 5) is involved in the synthesis of ceramide, a lipid molecule involved in several cellular signaling pathways. *Cers5* knockdown has been shown to reduce cardiomyocyte hypertrophy in mouse models<sup>29</sup>. However, it is unclear whether the blood pressure raising effects at this locus are the cause or result of any potential effects on cardiac hypertrophy. Future studies investigating this locus in relation to parameters of cardiac hypertrophy and function (*e.g.* ventricular wall thickness) should help address this question.



The DBP raising allele of the nsSNV (rs867186-A) was associated with increased expression of *PROCR* in adipose tissue ( $P_{\text{MuTHER}}=3.24 \times 10^{-15}$ ) and skin ( $P_{\text{MuTHER}}=1.01 \times 10^{-11}$ ) (Supplementary Figure 5). There was no evidence against colocalisation of the eQTL and DBP association thus supporting *PROCR* as a candidate causal gene. *PROCR* encodes the Endothelial Protein C receptor, a serine protease involved in the blood coagulation pathway, and rs867186 has previously been associated with coagulation and haematological factors.<sup>30,31</sup> The PP decreasing allele of, rs10407022-T, which is predicted to have detrimental effects on protein structure (Methods) was associated with increased expression of *AMH* in muscle ( $P_{\text{GTEx}}=9.95 \times 10^{-15}$ ), thyroid ( $P_{\text{GTEx}}=8.54 \times 10^{-7}$ ), nerve ( $P_{\text{GTEx}}=7.15 \times 10^{-8}$ ), tibial artery ( $P_{\text{GTEx}}=6.46 \times 10^{-9}$ ), adipose ( $P_{\text{GTEx}}=4.69 \times 10^{-7}$ ), and skin ( $P_{\text{GTEx}}=5.88 \times 10^{-8}$ ) (Supplementary Figure 5). There was no evidence against colocalisation of the eQTL and PP association, which supports *AMH* as a candidate causal gene for PP. Low AMH levels have been previously associated with hypertensive status in women with the protein acting as a marker of ovarian reserve<sup>32</sup>. The intergenic SBP raising allele of rs4728142-A was associated with reduced expression of *IRF5* in skin ( $P_{\text{MuTHER}}=5.24 \times 10^{-31}$ ) and LCLs ( $P_{\text{MuTHER}}=1.39 \times 10^{-34}$ ), whole blood ( $P_{\text{GTEx}}=3.12 \times 10^{-7}$ ) and tibial artery ( $P_{\text{GTEx}}=1.71 \times 10^{-7}$ ).

Three novel rare nsSNVs were identified that map to *RBM47*, *RRAS* (both associated with SBP) and *COL21A1* (associated with PP). They had larger effect sizes than common variant associations ( $>1.5$  mmHg per allele; Supplementary Figure 6) and were predicted to have detrimental effects on protein structure (Supplementary Table 16; Methods). In *RBM47*, rs35529250 (p.Gly538Arg) is located in a highly conserved region of the gene and was most strongly associated with SBP (MAF=0.008; +1.59 mmHg per T allele;  $P=5.90 \times 10^{-9}$ ). *RBM47* encodes the RNA binding motif protein 47 and is responsible for post-transcriptional regulation of RNA, through its direct and selective binding with the molecule.<sup>33</sup> In *RRAS*, rs61760904 (p.Asp133Asn) was most strongly associated with SBP (MAF=0.007; +1.51 mmHg per T allele;  $P=8.45 \times 10^{-8}$ ). *RRAS* encodes a small GTPase belonging to the Ras subfamily of proteins H-RAS, N-RAS, and K-RAS and has been implicated in actin cytoskeleton remodelling, and controlling cell proliferation, migration and cycle processes<sup>34</sup>. The nsSNV in *COL21A1* (rs200999181, p.Gly665Val) was most strongly associated with PP (MAF=0.001; +3.14 mmHg per A allele;  $P=1.93 \times 10^{-9}$ ). *COL21A1* encodes the collagen alpha-1 chain precursor of type XXI collagen, a member of the FACIT (fibril-associated collagens with an interrupted triple helix) family of proteins<sup>35</sup>. The gene is detected in many tissues, including the heart and aorta. Based on our results, these three genes represent good candidates for functional follow-up. However, due to the incomplete coverage of all SNVs across the region on the Exome chip, it is possible that other non-genotyped SNVs may better explain some of these associations. We therefore checked for variants in LD ( $r^2 > 0.3$ ) with these three rare nsSNVs in the UK10K + 1000G dataset<sup>36</sup> to ascertain if there are other candidate SNVs at these loci (Supplementary Table 17). There were no SNVs within 1Mb of the *RBM47* locus in LD with the BP associated SNV. At the *COL21A1* locus there were only SNVs in moderate LD, and these were annotated as intronic, intergenic or in the 5'UTR. At the *RRAS* locus, there were two SNVs in strong LD with the BP associated SNV, which both mapped to introns of *SCAF1* and are not predicted to be damaging. All SNVs in LD at both loci were rare as expected

(Supplementary Table 17) supporting a role for rare variants. Hence, the rare BP associated nsSNVs at *RBM47*, *COL21A1* and *RRAS* remain the best causal candidates.

### Pathway and network analyses

To identify connected gene sets and pathways implicated by the BP associated genes we used Meta-Analysis Gene-set Enrichment of variant Associations (MAGENTA)<sup>37</sup> and GeneGo MetaCore (Thomson Reuters, UK). MAGENTA tests for over-representation of BP associated genes in pre-annotated pathways (gene sets) (Methods and Supplementary Table 18a). GeneGo Metacore identifies potential gene networks. The MAGENTA analysis was used for hypothesis generation and results were compared with the GeneGo Metacore outputs to cross-validate findings.

Using MAGENTA there was an enrichment ( $P < 0.01$  and  $FDR < 5\%$  in either the EUR\_SAS or the EUR participants) of six gene sets with DBP, three gene sets with HTN and two gene sets for SBP (Supplementary Table 18b). The RNA polymerase I promoter clearance (chromatin modification) pathway showed the most evidence of enrichment with genes associated with DBP ( $P_{\text{Reactome}} = 8.4 \times 10^{-5}$ ,  $FDR = 2.48\%$ ). NOTCH signalling was the most associated pathway with SBP ( $P_{\text{Reactome}} = 3.00 \times 10^{-4}$ ,  $FDR = 5\%$ ) driven by associations at the *FURIN* gene. The inorganic cation anion solute carrier (SLC) transporter pathway had the most evidence of enrichment by HTN associated genes ( $P_{\text{Reactome}} = 8.00 \times 10^{-6}$ ,  $FDR = 2.13\%$ ).

Using GeneGo MetaCore, five network processes were enriched ( $FDR < 5\%$ ; Methods; Supplementary Tables 19 and 20). These included several networks with genes known to influence vascular tone and BP: inflammation signalling,  $P = 1.14 \times 10^{-4}$  and blood vessel development  $P = 2.34 \times 10^{-4}$ . The transcription and chromatin modification network ( $P = 2.85 \times 10^{-4}$ ) was also enriched, a pathway that was also highlighted in the MAGENTA analysis, with overlap of the same histone genes (*HIST1H4C*, *HIST1H2AC*, *HIST1H2BC*, *HIST1H1T*) and has also been recently reported in an integrative network analysis of published BP loci and whole blood expression profiling<sup>38</sup>. Two cardiac development pathways were enriched: the oxidative stress-driven (ROS/NADPH) ( $P = 4.12 \times 10^{-4}$ ) and the Wnt/ $\beta$ -catenin/integrin-driven ( $P = 0.0010$ ). Both these cardiac development pathways include the *MYH6*, *MYH7*, and *TBX2* genes, revealing a potential overlap with cardiomyopathies and hypertension, and suggesting some similarity in the underlying biological mechanisms.

### Discussion

By conducting the largest ever genetic study of BP, we identified further novel common variants with small effects on BP traits, similar to what has been observed for obesity and height<sup>39,40</sup>. More importantly, our study identified some of the first rare coding variants of strong effect ( $> 1.5 \text{ mmHg}$ ) that are robustly associated with BP traits in the general population, complementing and extending the previous discovery and characterisation of variants underlying rare Mendelian disorders of blood pressure regulation<sup>41</sup>. Using SNV associations in 17 genes reported to be associated with monogenic disorders of blood pressure (Methods) we found no convincing evidence of enrichment ( $P_{\text{enrichment}} = 0.044$ ).

This suggests that BP control in the general population may occur through different pathways to monogenic disorders of BP re-enforcing the importance of our study findings. The identification of 30 novel BP loci plus further new independent secondary signals within four novel and five known loci (Methods) has augmented the trait variance explained by 1.3%, 1.2% and 0.93% for SBP, DBP and PP respectively within our data-set. This suggests that with substantially larger sample sizes, for example through UK BioBank42, we expect to identify 1000s more loci associated with BP traits, and replicate more of our discovery SNV associations that are not yet validated in the current report.

The discovery of rare missense variants has implicated several interesting candidate genes, which are often difficult to identify from common variant GWAS, and should therefore lead to more rapidly actionable biology. *A2ML1*, *COL21A1*, *RRAS* and *RBM47* all warrant further follow-up studies to define the role of these genes in regulation of BP traits, as well as functional studies to understand their mechanisms of action. *COL21A1* and *RRAS* warrant particular interest since both are involved in blood vessel remodelling, a pathway of known aetiological relevance to hypertension.

We observed a rare nonsense SBP associated variant in *ENPEP* (rs33966350; p.Trp317\*): this overlaps a highly conserved region of both the gene and protein and is predicted to result in either a truncated protein with reduced catalytic function or is subject to nonsense mediated RNA decay. ENPEP converts angiotensin II (AngII) to Ang-III. AngII activates the angiotensin 1 (AT1) receptor resulting in vasoconstriction, while AngIII activates the angiotensin 2 (AT2) receptor that promotes vasodilation and protects against hypertension. 43 The predicted truncated protein may lead to predominant AngII signaling in the body, and increases in BP. This new observation could potentially inform therapeutic strategies. Of note, angiotensin-converting-enzyme (ACE) inhibitors are commonly used in the treatment of hypertension. However, patients who suffer from adverse reactions to ACE inhibitors, such as dry cough and skin rash, would benefit from alternative drugs that target RAAS. Murine studies have shown that in the brain, AngIII is the preferred AT1 agonist that promotes vasoconstriction and increases blood pressure, as opposed to AngII in the peripheral system. These results have motivated the development of brain specific APA inhibitors to treat hypertension44. Our results confirm APAs, such as ENPEP, as a valid target to modify blood pressure, but suggest that long-term systemic reduction in APA activity may lead to an increase in blood pressure. Future studies are needed to examine the effects of the p.Trp317\* variant on the RAAS system, specifically in the brain and peripheral vasculature, in order to test the benefits of the proposed therapeutic strategy in humans.

In addition to highlighting new genes in pathways of established relevance to BP and hypertension, and identifying new pathways, we have also identified multiple signals at new loci. For example, there are three distinct signals at the locus containing the *MYH6/MYH7* genes, and we note that *TBX2* maps to one of the novel regions. These genes are related to cardiac development and/or cardiomyopathies, and provide an insight into the shared inheritance of multiple complex traits. Unravelling the causal networks within these polygenic pathways may provide opportunities for novel therapies to treat or prevent both hypertension and cardiomyopathies.

## Online Methods

### Overview of discovery studies

The cohorts contributing to the discovery meta-analyses comprise studies from three consortia (CHD Exome+, ExomeBP, and GoT2D/T2D-GENES) with a total number of 192,763 unique samples. All participants provided written informed consent and the studies were approved by their local Research Ethics Committees and/or Institutional Review Boards.

The CHD Exome+ consortium comprised 77,385 samples: eight studies (49,898 samples) of European (EUR) ancestry, two studies (27,487 samples) of South Asian (SAS) ancestry (Supplementary Table 1). The ExomeBP consortium included 25 studies (75,620 samples) of EUR ancestry (Supplementary Table 1). The GoT2D consortium comprised 14 studies (39,758 samples) of Northern EUR ancestry from Denmark, Finland, and Sweden (Supplementary Table 1). The participating studies and their characteristics including BP phenotypes are detailed in Supplementary Tables 1 and 2. Note, any studies contributing to multiple consortia were only included once in all meta-analyses.

### Phenotypes

Four blood pressure (BP) traits were analysed: systolic blood pressure (SBP), diastolic blood pressure (DBP), pulse pressure (PP) and hypertension (HTN). For individuals known to be taking BP lowering medication, 15/10 mmHg was added to the raw SBP/DBP values, respectively, to obtain medication-adjusted SBP/DBP values<sup>45</sup>. PP was defined as SBP minus DBP, post-adjustment. For HTN, individuals were classified as hypertensive cases if they satisfied at least one of: (i) SBP  $\geq$  140 mmHg, (ii) DBP  $\geq$  90 mmHg, (iii) taking anti-hypertensive or BP lowering medication. All other individuals were included as controls. The four BP traits were correlated (SBP:DBP correlations were between 0.6 and 0.8, and SBP:PP correlations were  $\sim$ 0.8). However, they measure partly distinct physiological features including, cardiac output, vascular resistance, and arterial stiffness, all measures for determining a cardiovascular risk profile. Therefore the genetic architecture of the individual phenotypes are of interest, and a multi-phenotype mapping approach was not adopted.

### Genotyping

All samples were genotyped using one of the Illumina HumanExome Beadchip arrays (Supplementary Table 3). An Exome chip quality control Standard Operating Procedure (SOP) developed by Anubha Mahajan, Neil Robertson and Will Rayner at the Wellcome Trust Centre for Human Genetics, University of Oxford was used by most studies for genotype calling and QC<sup>46</sup> (Supplementary Table 3). All genotypes were aligned to the plus strand of the human genome reference sequence (Build37) prior to any analyses and any unresolved mappings were removed. Genotype cluster plots were reviewed for all the novel rare variants (both lead and secondary signals) and for rare variants that contributed to the gene-based testing.

## Meta-analyses

Meta-analyses were performed using METAL47, for both discovery and replication analyses, using inverse variance weighted fixed effect meta-analysis for the continuous traits (SBP, DBP and PP) and sample size weighted meta-analysis for the binary trait (HTN).

## Discovery SNV analyses

Analyses of both untransformed and inverse normal transformed SBP, DBP and PP were conducted within each contributing study. The analyses of the transformed traits were performed in order to minimise sensitivity to deviations from normality in the analysis of rare variants and for discovery of new SNV-BP associations. The residuals from the null model obtained after regressing the medication-adjusted trait on the covariates (age, age<sup>2</sup>, sex, BMI, and disease status for CHD) within a linear regression model, were ranked and inverse normalised. These normalised residuals were used to test trait-SNV associations. All SNVs that passed QC were analysed for association, without any further filtering by MAF, but a minor allele count of 10 was used for the analysis of HTN. An additive allelic effects model was assumed.

Two meta-analyses were performed for each trait, one with EUR and SAS ancestries combined (EUR\_SAS) and another for EUR ancestry alone. Contributing studies used principal components (PCs) to adjust for population stratification. Consequently minimal inflation in the association test statistics,  $\lambda$ , was observed ( $\lambda=1.07$  for SBP, 1.10 for DBP, 1.04 for PP and  $<1$  for HTN in the transformed discovery meta-analysis in EUR\_SAS;  $\lambda=1.06$  for SBP, 1.09 for DBP, 1.05 for PP and  $<1$  for HTN in the transformed discovery meta-analysis in EUR; Supplementary Figure 7). The meta-analyses were performed independently in two centres and results were found to be concordant between centres. Given the studies contributing to the discovery analyses were ascertained on CHD or T2D, we tested potential systematic bias in calculated effect estimates amongst these studies. No evidence of bias in the overall effect estimates was obtained.

The results for the transformed traits were taken forward and used to select candidate SNVs for replication. Results ( $P$ -values) from the transformed and untransformed analyses were strongly correlated ( $r^2>0.9$ ).

## Replication SNV analyses

SNVs associated with any of the transformed traits (SBP, DBP, PP) or HTN were annotated using the Illumina SNV annotation file, humanexome-12v1\_a\_gene\_annotation.txt, independently across two centres. Given the difference in power to detect common versus low frequency and rare variant associations, two different significance thresholds were chosen for SNV selection. For SNVs with MAF  $>0.05$ ,  $P < 1 \times 10^{-5}$  was selected, while,  $P < 1 \times 10^{-4}$  was used for SNVs with MAF  $< 0.05$ . By choosing a significance threshold of  $P < 1 \times 10^{-4}$  we maximized the opportunity to follow-up rare variants (making the assumption that any true signals at this threshold could replicate at Bonferroni adjusted significance,  $P < 6.17 \times 10^{-4}$ , assuming  $\alpha=0.05$  for 81 SNVs). All previously published BP associated SNVs and any variants in LD with them ( $r^2>0.2$ ), were removed from the list of associated SNVs as we aimed to replicate new findings only. SNVs for which only one study contributed to

the association result or showed evidence of heterogeneity ( $P_{\text{het}} < 0.0001$ ) were removed from the list as they were likely to be an artefact. Where SNVs were associated with multiple traits, to minimise the number of tests performed, only the trait with the smallest  $P$ -value was selected as the primary trait in which replication was sought. Where multiple SNVs fitted these selection criteria for a single region, only the SNV with the smallest  $P$ -value was selected. In total, 81 SNVs were selected for validation in independent samples. These 81 SNVs had concordant association results for both transformed and non-transformed traits. Eighty SNVs were selected from EUR\_SAS results (with consistent support in EUR), and one SNV from EUR results only. In the next step, we looked up the 81 SNV-BP associations using data from a separate consortium, the CHARGE+ exome chip blood pressure consortium (who had analysed untransformed SBP, DBP, PP and HTN), and UHP and Lolipop (ExomeBP consortium; Supplementary Tables 2 and 3). The analysed residuals from CHARGE+ were approximately normally distributed in their largest studies (Supplementary Figure 8).

Two meta-analyses of the replication datasets were performed: one of EUR samples, and a second of EUR, African American, Hispanics and SAS ancestries ("ALL"). Replication was confirmed if  $P(1\text{-tailed}) < 0.05/81 = 6.17 \times 10^{-4}$  and the effect (beta) was in the direction observed in discovery meta-analyses for the selected trait. A combined meta-analysis was performed of discovery (untransformed results as only untransformed data was available from CHARGE+ exome chip blood pressure consortium) and replication results across the four traits to assess the overall support for each locus. For the combined meta-analyses, a GWS threshold of,  $P \ 5 \times 10^{-8}$ , was used to declare a SNV as novel rather than a less stringent experiment wide threshold, as GWS is used to declare significance in GWAS and we wish to minimise the possibility of false positive associations. (Note that GWS is equivalent to an exome-wide threshold of  $P \ 2 \times 10^{-7}$  adjusted for four traits).

Note: all validated BP-associated variants were associated at  $P < 10^{-5}$  in the discovery dataset (for the primary trait). Hence, we could have used the same inclusion criteria for both common and rare SNVs. Therefore the optimal threshold to choose for future experiments may need further consideration.

### Conditional analyses and gene-based tests

The RAREMETALWORKER (RMW) tool<sup>15</sup> (version 4.13.3) that does not require individual level data to perform conditional analyses and gene-based tests was used for conditional analyses. All studies that contributed to the SNV discovery analyses were re-contacted and asked to run RMW. Only FENLAND, GoDARTS, HELIC-MANOLIS, UKHLS and EPIC-InterAct were unable to run RMW, while two new studies were included, INCIPE and NFBC1966 (Supplementary Table 1 and 2). In total, 43 studies (147,402 samples) were included in the EUR analyses and 45 studies (173,329 samples) in the EUR\_SAS analyses (Supplementary Tables 2 and 3). Comparison of discovery and RMW study level results were made (Supplementary Information).

For each novel locus, the genomic coordinates and size of the region were defined according to recombination rates (Supplementary Table 9) around the lead variant. For known loci, a 1 Mb window was used (Supplementary Table 14). Conditional analyses were performed



across each region, in both EUR and EUR\_SAS samples, for the transformed phenotype corresponding to the validated BP trait for novel loci and the published BP trait for known loci.

Gene based tests were performed in both the EUR and EUR\_SAS datasets using the Sequence Kernel Association Test (SKAT)<sup>16</sup> method implemented in RMW as it allows for the SNVs to have different directions and magnitudes of effect. Burden tests were also performed but are not presented as only SKAT provided significant results. The variants in the gene-based tests using SKAT were weighted using the default settings, *i.e.* a beta distribution density function to up-weight rare variants,  $\text{Beta}(\text{MAF}_j, 1, 25)$  where  $\text{MAF}_j$  represents the pooled MAF for variant  $j$  across all studies. Analyses were restricted to coding SNVs with  $\text{MAF} < 5\%$  and  $< 1\%$ . Genes were deemed to be associated if  $P < 2.8 \times 10^{-6}$  (Bonferroni adjusted for 17,996 genes). To confirm the gene associations were not attributable to a solitary SNV, a gene-based test conditional on the most associated SNV was performed ( $P_{\text{conditional}} < 0.001$ ). The QC of all SNVs contributing to the gene based tests including the number of samples and studies were checked prior to claiming association. We sought replication of associated genes in the CHARGE+ exome chip blood pressure consortium.

### Pathway analyses with MAGENTA

We tested seven databases in MAGENTA<sup>37</sup> (BioCarta, Kyoto Encyclopedia of Genes and Genomes, Ingenuity, Panther, Panther Biological Processes, Panther Molecular Functions and Reactome) for overrepresentation of the SNV discovery results from both EUR and EUR\_SAS ancestries. Each of the four BP phenotypes were tested. Pathways exhibiting  $P < 0.01$  and  $\text{FDR} < 5\%$  were considered statistically significant.

### GeneGo MetaCore Network analyses

A set of BP genes based on previously published studies and our current results (locus defined as  $r^2 > 0.4$  and 500kb on either side of the lead SNV; Supplementary Table 19) were tested for enrichment using the THOMSON REUTERS MetaCore™ Single Experiment Analysis workflow tool. The data were mapped onto selected MetaCore ontology databases: pathway maps, process networks, GO processes and diseases / biomarkers, for which functional information is derived from experimental literature. Outputs were sorted based on  $P$ - and  $\text{FDR}$ -values. A gene set was considered enriched for a particular process if  $P < 0.05$  and  $\text{FDR} < 5\%$ .

### Genetic Risk Score

To assess the effect of BP on CHD, ischemic stroke (and subtypes: large vessel, small vessel and cardioembolic stroke) left ventricular mass, left ventricular wall thickness, heart failure, HDL-c, LDL-c, total cholesterol, triglycerides and eGFR, we performed a weighted generalized linear regression of the genetic associations with each outcome variable on the genetic associations with BP.

When genetic variants are uncorrelated, the estimates from such a weighted linear regression analysis using summarized data, and a genetic risk score analysis using individual-level data,

are equal<sup>48</sup>. We refer to the analysis as a genetic risk score (also known as a polygenic risk score) analysis as this is likely to be more familiar to applied readers. As some of the genetic variants in our analysis are correlated, a generalized weighted linear regression model is fitted that accounts for the correlations between variants, as follows: If  $\beta_X$  are the genetic associations (beta-coefficients) with the risk factor (here, BP) and  $\beta_Y$  are the genetic associations with the outcome, then the causal estimate from a weighted generalized linear regression is  $(\beta_X^T \Omega^{-1} \beta_X)^{-1} \beta_X^T \Omega^{-1} \beta_Y$ , with standard error,

$$\hat{\sigma} \sqrt{(\beta_X^T \Omega^{-1} \beta_X)^{-1}},$$

where  $^T$  is a matrix transpose,  $\hat{\sigma}$  is the estimate of the residual standard error from the regression model, and the weighting matrix  $\Omega$  has terms

$$\Omega_{j_1 j_2} = \sigma_{Y_{j_1}} \sigma_{Y_{j_2}} \rho_{j_1 j_2}$$

, where  $\sigma_{Y_j}$  is the standard error of the genetic association with the outcome for the  $j$ th SNV, and  $\rho_{j_1 j_2}$  is the correlation between the  $j_1$ th and  $j_2$ th SNVs. The presence of the estimated residual standard error allows for heterogeneity between the causal estimates from the individual SNVs as overdispersion in the regression model (in the case of underdispersion, the residual standard error estimate is set to unity). This is equivalent to combining the causal estimates from each SNV using a multiplicative random-effects model<sup>49</sup>.

For each of SBP, DBP and PP, the score was created using both the novel and known BP SNVs or a close proxy ( $r^2 > 0.8$ ). Both the sentinel SNV association and any secondary SNV associations that remained after adjusting for the sentinel SNV were included in the genetic risk score. For the 30 validated novel SNV-BP associations,  $\beta$ s were taken from the independent replication analyses (Table 1 and 2) to weight the SNV in the genetic risk score. For the secondary SNVs from the seven novel loci and five known loci,  $\beta$ s were taken from the discovery analyses (Supplementary Tables 10 and 15). For the 82 known SNVs, 43 were either genotyped or had proxies on the Exome chip and the  $\beta$ s were taken from discovery results (Supplementary Table 13), the remaining  $\beta$ s were taken from published effect estimates. This strategy for selecting betas for use in the GRS was taken to minimize the influence of winner's curse. The associations between the BP variants with CHD, HDL-c, LDL-c, total cholesterol, log(triglycerides) and log(eGFR) were obtained using the CHD Exome+ Consortium studies, the associations with BMI, waist-hip ratio adjusted BMI and height from the GIANT consortium (unpublished data), ischemic stroke from METASTROKE<sup>25</sup>, and left ventricular mass, left ventricular wall thickness and heart failure from EchoGen<sup>27</sup> and CHARGE-HF<sup>26</sup>. A causal interpretation of the association of GRS with the outcome as the effect of BP on the outcome assumes that the effects of genetic variants on the outcome are mediated via blood pressure and not via alternate causal pathways, for example via LV thickness. There are also limitations of the Mendelian randomization approach in distinguishing between the causal effects of different measures of blood pressure, due to the paucity of genetic variants associated with only one measure of blood pressure.

## eQTL analyses

The MuTHER dataset contains gene expression data from 850 UK twins for 23,596 probes and 2,029,988 (HapMap 2 imputed) SNVs. All cis-associated SNVs with  $FDR < 1\%$ , within each of the 30 novel regions (IMPUTE info score  $> 0.8$ ) were extracted from the MuTHER project dataset for, LCL ( $n=777$ ), adipose ( $n=776$ ) and skin ( $n=667$ ) 50. The pilot phase of the GTEx Project (dbGaP Accession phs000424.v3.p1) provides expression data from up to 156 individuals for 52,576 genes and 6,820,472 genotyped SNVs (imputed to 1000 Genomes project, MAF  $\geq 5\%$ ) 51. The eQTL analysis was focused on subcutaneous adipose tissue ( $n=94$ ), tibial artery ( $n=112$ ), heart (left ventricle) ( $n=83$ ), lung ( $n=119$ ), skeletal muscle ( $n=138$ ), tibial nerve ( $n=88$ ), skin (sun exposed, lower leg) ( $n=96$ ), thyroid ( $n=105$ ) and whole blood ( $n=156$ ) which have  $> 80$  samples and genes expressed at least 0.1 RPKM in 10 or more individuals in a given tissue. All transcripts with a transcription start site (TSS) within one of the 30 new BP loci and for which there was a cis-associated SNV (IMPUTE info score  $> 0.4$ ) within 1Mb of the TSS at  $FDR < 5\%$ , were identified. Kidney was not evaluated because the sample size was too small ( $n=8$ ). From each resource, we report eQTL signals, which reach the resource-specific thresholds for significance described above, for SNVs that are in LD ( $r^2 > 0.8$ ) with our sentinel SNV.

For identified eQTLs, we tested whether they colocalised with the BP associated SNV 52. Colocalisation analyses were considered to be significant if the posterior probability of colocalisation was greater than 0.95.

## Annotation of variants

*In silico* prediction of the functional effect of associated variants was based on the annotation from dbSNP, the Ensembl Variant Effect Predictor tool and the Exome Variant Server, NHLBI GO Exome Sequencing Project (ESP), Seattle, WA.

## Trait variance explained

The percentage trait variance explained for SBP, DBP, PP was assessed with 5,861 individuals with complete information for all phenotypes and covariates from the population-based cohort, 1958BC.

Two genetic models were investigated: one containing the 43 previously known BP associated SNVs covered on the Exome chip; the other additionally including the 30 novel lead SNVs and 9 conditionally independent SNVs from both novel and known loci. These nine conditionally independent SNVs were taken from the EUR results, as 1958BC is EUR. They included four from novel loci (*PREX1*, *COL21A1*, *PRKAG1* and *MYH6* (there was only 1 in EUR); Supplementary Table 10) and five from known loci (*ST7L-CAPZA1-MOV10*, *FIGN-GRB14*, *ENPEP*, *TBX5-TBX3* and *HOXC4*; Supplementary Table 15).

The residual trait was obtained by adjusting each of the BP traits in a regression model with sex and BMI variables (not age or age<sup>2</sup> as all 1958BC individuals were aged 44 years). The residual trait was regressed on all SNVs within the corresponding model and adjusted for the first ten PCs. The  $R^2$  calculated from this regression model was used as the percentage trait variance explained.

## Monogenic Enrichment analyses

To determine if sub-significant signals of association were present in a set of genes associated with monogenic forms of disease, we performed an enrichment analysis of the discovery single variant meta-analyses association results for all four traits, both for EUR and EUR\_SAS datasets.

The monogenic gene set included: *WNK1*, *WNK4*, *KLHL3*, *CUL3*, *PPARG*, *NR3C2*, *CYP11B1*, *CYP11B2*, *CYP17A1*, *HSD11B2*, *SCNN1A*, *SCNN1B*, *SCNN1G*, *CLCNKB*, *KCNJ1*, *SLC12A1*, *SLC12A3*. The association results of coding SNVs in these genes were extracted and the number of tests with  $P < 0.001$  observed. In order to determine how often such an observation would be observed by chance, we constructed 1,000 matched gene sets. The matching criteria for each monogenic gene was the intersection of all genes in the same exon length quintile and all genes in the same coding variant count decile. Within the matched sets, the number of variants with  $P < 0.001$  was observed. The empirical  $P$ -value was calculated as the fraction of matched sets with an equal or larger number of variants less than 0.001.

## Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

## Authors

Praveen Surendran<sup>#1</sup>, Fotios Drenos<sup>#2,3</sup>, Robin Young<sup>#1</sup>, Helen Warren<sup>#4,5</sup>, James P Cook<sup>#6,7</sup>, Alisa K Manning<sup>#8,9,10</sup>, Niels Grarup<sup>#11</sup>, Xueling Sim<sup>#12,13,14</sup>, Daniel R Barnes<sup>1</sup>, Kate Witkowska<sup>4,5</sup>, James R Staley<sup>1</sup>, Vinicius Tragante<sup>15</sup>, Taru Tukiainen<sup>8,9,16</sup>, Hanieh Yaghootkar<sup>17</sup>, Nicholas Masca<sup>18,19</sup>, Daniel F Freitag<sup>1</sup>, Teresa Ferreira<sup>20</sup>, Olga Giannakopoulou<sup>21</sup>, Andrew Tinker<sup>21,5</sup>, Magdalena Harakalova<sup>15</sup>, Evelin Mihailov<sup>22</sup>, Chunyu Liu<sup>23</sup>, Aldi T Kraja<sup>24,25</sup>, Sune Fallgaard Nielsen<sup>26</sup>, Asif Rasheed<sup>27</sup>, Maria Samuel<sup>27</sup>, Wei Zhao<sup>28</sup>, Lori L Bonnycastle<sup>29</sup>, Anne U Jackson<sup>13,12</sup>, Narisu Narisu<sup>29</sup>, Amy J Swift<sup>29</sup>, Lorraine Southam<sup>30,20</sup>, Jonathan Marten<sup>31</sup>, Jeroen R Huyghe<sup>13,12</sup>, Alena Stan áková<sup>32</sup>, Cristiano Fava<sup>33,34</sup>, Therese Ohlsson<sup>33</sup>, Angela Matchan<sup>30</sup>, Kathleen E Stirrups<sup>21,35</sup>, Jette Bork-Jensen<sup>11</sup>, Anette P Gjesing<sup>11</sup>, Jukka Kontto<sup>36</sup>, Markus Perola<sup>36,37,22</sup>, Susan Shaw-Hawkins<sup>4</sup>, Aki S Havulinna<sup>36</sup>, He Zhang<sup>38</sup>, Louise A Donnelly<sup>39</sup>, Christopher J Groves<sup>40</sup>, N William Rayner<sup>40,20,30</sup>, Matt J Neville<sup>40,41</sup>, Neil R Robertson<sup>20,40</sup>, Andrianos M Yiorkas<sup>42,43</sup>, Karl-Heinz Herzig<sup>44,45</sup>, Eero Kajantie<sup>36,46,47</sup>, Weihua Zhang<sup>48,49</sup>, Sara M Willems<sup>50</sup>, Lars Lannfelt<sup>51</sup>, Giovanni Malerba<sup>52</sup>, Nicole Soranzo<sup>53,35,54</sup>, Elisabetta Trabetti<sup>52</sup>, Niek Verweij<sup>55,9,56</sup>, Evangelos Evangelou<sup>48,57</sup>, Alireza Moayyeri<sup>48,58</sup>, Anne-Claire Vergnaud<sup>48</sup>, Christopher P Nelson<sup>18,19</sup>, Alaitz Poveda<sup>59,60</sup>, Tibor V Varga<sup>59</sup>, Muriel Caslake<sup>61</sup>, Anton JM de Craen<sup>62,63</sup>, Stella Trompet<sup>62,64</sup>, Jian'an Luan<sup>50</sup>, Robert A Scott<sup>50</sup>, Sarah E Harris<sup>65,66</sup>, David CM Liewald<sup>65,67</sup>, Riccardo Marioni<sup>65,66,68</sup>, Cristina Menni<sup>69</sup>, Aliki-Eleni Farmaki<sup>70</sup>, Göran Hallmans<sup>71</sup>, Frida Renström<sup>59,71</sup>, Jennifer E Huffman<sup>31,23</sup>, Maija Hassinen<sup>72</sup>, Stephen Burgess<sup>1</sup>, Ramachandran S Vasan<sup>23,73,74</sup>, Janine F Felix<sup>75</sup>, CHARGE-Heart Failure Consortium<sup>76</sup>, Maria Uria-Nickelsen<sup>77</sup>, Anders

Malarstig<sup>78</sup>, Dermot F Reily<sup>79</sup>, Maarten Hoek<sup>80</sup>, Thomas Vogt<sup>80,81</sup>, Honghuang Lin<sup>23,82</sup>, Wolfgang Lieb<sup>83</sup>, EchoGen Consortium<sup>76</sup>, Matthew Traylor<sup>84</sup>, Hugh F Markus<sup>84</sup>, METASTROKE Consortium<sup>76</sup>, Heather M Highland<sup>85</sup>, Anne E Justice<sup>85</sup>, Eirini Marouli<sup>21</sup>, GIANT Consortium<sup>76</sup>, Jaana Lindström<sup>36</sup>, Matti Uusitupa<sup>86,87</sup>, Pirjo Komulainen<sup>72</sup>, Timo A Lakka<sup>72,88,89</sup>, Rainer Rauramaa<sup>72,89</sup>, Ozren Polasek<sup>90,91</sup>, Igor Rudan<sup>90</sup>, Olov Rolandsson<sup>92</sup>, Paul W Franks<sup>59,92,93</sup>, George Dedoussis<sup>70</sup>, Timothy D Spector<sup>69</sup>, EPIC-InterAct Consortium<sup>76</sup>, Pekka Jousilahti<sup>36</sup>, Satu Männistö<sup>36</sup>, Ian J Deary<sup>65,67</sup>, John M Starr<sup>65,94</sup>, Claudia Langenberg<sup>50</sup>, Nick J Wareham<sup>50</sup>, Morris J Brown<sup>4</sup>, Anna F Dominiczak<sup>95</sup>, John M Connell<sup>39</sup>, J Wouter Jukema<sup>64,96</sup>, Naveed Sattar<sup>95</sup>, Ian Ford<sup>61</sup>, Chris J Packard<sup>61</sup>, Tõnu Esko<sup>22,97,8,9</sup>, Reedik Mägi<sup>22</sup>, Andres Metspalu<sup>22,98</sup>, Rudolf A de Boer<sup>99</sup>, Peter van der Meer<sup>99</sup>, Pim van der Harst<sup>99,100,101</sup>, Lifelines Cohort Study<sup>76</sup>, Giovanni Gambaro<sup>102</sup>, Erik Ingelsson<sup>103,104</sup>, Lars Lind<sup>103</sup>, Paul IW de Bakker<sup>105,106</sup>, Mattijs E Numans<sup>107,106</sup>, Ivan Brandslund<sup>108,109</sup>, Cramer Christensen<sup>110</sup>, Eva RB Petersen<sup>111</sup>, Eeva Korpi-Hyövälti<sup>112</sup>, Heikki Oksa<sup>113</sup>, John C Chambers<sup>48,49,114</sup>, Jaspal S Kooner<sup>49,115,114</sup>, Alexandra IF Blakemore<sup>42,43</sup>, Steve Franks<sup>116</sup>, Marjo-Riitta Jarvelin<sup>117,118,119,120</sup>, Lise L Husemoen<sup>121</sup>, Allan Linneberg<sup>121,122,123</sup>, Tea Skaaby<sup>121</sup>, Betina Thuesen<sup>121</sup>, Fredrik Karpe<sup>40,41</sup>, Jaakko Tuomilehto<sup>36,124,125,126</sup>, Alex SF Doney<sup>39</sup>, Andrew D Morris<sup>127</sup>, Colin NA Palmer<sup>39</sup>, Oddgeir Lingaas Holmen<sup>128,129</sup>, Kristian Hveem<sup>128,130</sup>, Cristen J Willer<sup>38,131,132</sup>, Tiinamaija Tuomi<sup>133,134</sup>, Leif Groop<sup>135,134</sup>, AnneMari Käräjämäki<sup>136,137</sup>, Aarno Palotie<sup>16,9,134,138</sup>, Samuli Ripatti<sup>134,139,30</sup>, Veikko Salomaa<sup>36</sup>, Dewan S Alam<sup>140</sup>, Abdulla al Shafi Majumder<sup>141</sup>, Emanuele Di Angelantonio<sup>1,54</sup>, Rajiv Chowdhury<sup>1</sup>, Mark I McCarthy<sup>40,41,20</sup>, Neil Poulter<sup>142</sup>, Alice V Stanton<sup>143</sup>, Peter Sever<sup>142</sup>, Philippe Amouyel<sup>144,145,146,147</sup>, Dominique Arveiler<sup>148</sup>, Stefan Blankenberg<sup>149,150</sup>, Jean Ferrières<sup>151</sup>, Frank Kee<sup>152</sup>, Kari Kuulasmaa<sup>36</sup>, Martina Müller-Nurasyid<sup>153,154,155</sup>, Giovanni Veronesi<sup>156</sup>, Jarmo Virtamo<sup>36</sup>, Panos Deloukas<sup>21,157</sup>, Wellcome Trust Case Control Consortium<sup>76</sup>, Paul Elliott<sup>117</sup>, Understanding Society Scientific Group<sup>76</sup>, Eleftheria Zeggini<sup>30</sup>, Sekar Kathiresan<sup>56,158,159,9</sup>, Olle Melander<sup>33</sup>, Johanna Kuusisto<sup>32</sup>, Markku Laakso<sup>32</sup>, Sandosh Padmanabhan<sup>95</sup>, David Porteous<sup>66</sup>, Caroline Hayward<sup>31</sup>, Generation Scotland<sup>160</sup>, Francis S Collins<sup>29</sup>, Karen L Mohlke<sup>161</sup>, Torben Hansen<sup>11</sup>, Oluf Pedersen<sup>11</sup>, Michael Boehnke<sup>13,12</sup>, Heather M Stringham<sup>13,12</sup>, EPIC-CVD Consortium<sup>76</sup>, Philippe Frossard<sup>27</sup>, Christopher Newton-Cheh<sup>56,158</sup>, CHARGE+ Exome Chip Blood Pressure Consortium<sup>76</sup>, Martin D Tobin<sup>6</sup>, Børge Grønne Nordestgaard<sup>26</sup>, T2D-GENES Consortium<sup>76</sup>, GoT2DGenes Consortium<sup>76</sup>, ExomeBP Consortium<sup>76</sup>, CHD Exome+ Consortium<sup>76</sup>, Mark J Caulfield<sup>4,5</sup>, Anubha Mahajan<sup>20</sup>, Andrew P Morris<sup>20,7</sup>, Maciej Tomaszewski<sup>18,19,162</sup>, Nilesh J Samani<sup>18,19</sup>, Danish Saleheen<sup>#28,27,1</sup>, Folkert W Asselbergs<sup>#15,101,163</sup>, Cecilia M Lindgren<sup>#164,9,20</sup>, John Danesh<sup>#1,165,54</sup>, Louise V Wain<sup>#6</sup>, Adam S Butterworth<sup>#1,166</sup>, Joanna MM Howson<sup>#1,168</sup>, and Patricia B Munroe<sup>#4,5,168</sup>

## Affiliations

<sup>1</sup>Cardiovascular Epidemiology Unit, Department of Public Health and Primary Care, University of Cambridge, Cambridge, UK <sup>2</sup>Medical Research Council Integrative Epidemiology Unit, School of Social and Community Medicine, University of Bristol,

Oakfield House, Oakfield Grove, Bristol, UK <sup>3</sup>Centre for Cardiovascular Genetics, Institute of Cardiovascular Science, Rayne Building University College London, London, UK <sup>4</sup>Clinical Pharmacology, William Harvey Research Institute, Queen Mary University of London, London, UK <sup>5</sup>National Institute for Health Research Barts Cardiovascular Biomedical Research Unit, Queen Mary University of London, London, UK <sup>6</sup>Department of Health Sciences, University of Leicester, Leicester, UK <sup>7</sup>Department of Biostatistics, University of Liverpool, Liverpool, UK <sup>8</sup>Department of Genetics, Harvard Medical School, Boston, Massachusetts, USA <sup>9</sup>Program in Medical and Population Genetics, Broad Institute, 7 Cambridge Center, Cambridge, Massachusetts, USA <sup>10</sup>Department of Molecular Biology, Massachusetts General Hospital, Boston, Massachusetts, USA <sup>11</sup>The Novo Nordisk Foundation Center for Basic Metabolic Research, Faculty of Health and Medical Sciences, University of Copenhagen, Copenhagen, Denmark <sup>12</sup>Center for Statistical Genetics, University of Michigan, Ann Arbor, Michigan, USA <sup>13</sup>Department of Biostatistics, University of Michigan, Ann Arbor, Michigan, USA <sup>14</sup>Saw Swee Hock School of Public Health, National University of Singapore, National University Health System, Singapore <sup>15</sup>Department of Cardiology, University Medical Center Utrecht, Utrecht, The Netherlands <sup>16</sup>Analytic and Translational Genetics Unit, Department of Medicine, Massachusetts General Hospital, Boston, Massachusetts, USA <sup>17</sup>Genetics of Complex Traits, Institute of Biomedical and Clinical Science, University of Exeter Medical School, Exeter, UK <sup>18</sup>Department of Cardiovascular Sciences, University of Leicester, Leicester, UK <sup>19</sup>National Institute for Health Research Leicester Biomedical Research Unit in Cardiovascular Disease, Leicester, UK <sup>20</sup>Wellcome Trust Centre for Human Genetics, Nuffield Department of Medicine, University of Oxford, Oxford, UK <sup>21</sup>Heart Centre, William Harvey Research Institute, Barts and The London School of Medicine and Dentistry, Queen Mary University of London, London, UK <sup>22</sup>Estonian Genome Center, University of Tartu, Tartu, Estonia <sup>23</sup>National Heart, Lung, and Blood Institute's and Boston University's Framingham Heart Study, Framingham, Massachusetts, USA <sup>24</sup>Division of Statistical Genomics, Center for Genome Sciences and Systems Biology, Washington University School of Medicine, St. Louis, Missouri, USA <sup>25</sup>Department of Genetics, Washington University School of Medicine, St. Louis, Missouri, USA <sup>26</sup>Department of Clinical Biochemistry Herlev Hospital, Copenhagen University Hospital, Herlev, Denmark <sup>27</sup>Centre for Non-Communicable Diseases, Karachi, Pakistan <sup>28</sup>Department of Biostatistics and Epidemiology, Perelman School of Medicine, University of Pennsylvania, Philadelphia, Pennsylvania, USA <sup>29</sup>Medical Genomics and Metabolic Genetics Branch, National Human Genome Research Institute, NIH, Bethesda, Maryland, USA <sup>30</sup>Wellcome Trust Sanger Institute, Genome Campus, Hinxton, UK <sup>31</sup>Medical Research Council Human Genetics Unit, Medical Research Council Institute of Genetics and Molecular Medicine, University of Edinburgh, Edinburgh, UK <sup>32</sup>Department of Medicine, University of Eastern Finland and Kuopio University Hospital, Kuopio, Finland <sup>33</sup>University of Lund, Department of Clinical Sciences, Malmö, Sweden <sup>34</sup>University of Verona, Department of Medicine, Verona, Italy <sup>35</sup>Department of Haematology, University of Cambridge, Cambridge, UK



<sup>36</sup>Department of Health, National Institute for Health and Welfare, Helsinki, Finland  
<sup>37</sup>Institute of Molecular Medicine FIMM, University of Helsinki, Finland <sup>38</sup>Department of Internal Medicine, Division of Cardiovascular Medicine, University of Michigan, Ann Arbor, Michigan, USA <sup>39</sup>Medical Research Institute, University of Dundee, Ninewells Hospital and Medical School, Dundee, UK <sup>40</sup>Oxford Centre for Diabetes, Endocrinology, and Metabolism, Radcliffe Department of Medicine, University of Oxford, Oxford, UK <sup>41</sup>National Institute for Health Research Oxford Biomedical Research Centre, Oxford University Hospital Trusts, Oxford, UK <sup>42</sup>Section of Investigative Medicine, Imperial College London, London, UK <sup>43</sup>Department of Life Sciences, Brunel University London, London, UK <sup>44</sup>Institute of Biomedicine, Biocenter Oulu, University of Oulu, Oulu, Finland <sup>45</sup>Department of Gastroenterology and Metabolism, Poznan University of Medical Sciences, Poznan, Poland <sup>46</sup>Hospital for Children and Adolescents, Helsinki University Central Hospital and University of Helsinki, Helsinki, Finland <sup>47</sup>Department of Obstetrics and Gynaecology, Oulu University Hospital and University of Oulu, Oulu, Finland <sup>48</sup>Department of Epidemiology and Biostatistics, School of Public Health, Imperial College London, London, UK <sup>49</sup>Department of Cardiology, Ealing Hospital, Middlesex, UK <sup>50</sup>Medical Research Council Epidemiology Unit, University of Cambridge School of Clinical Medicine, Box 285 Institute of Metabolic Science, Cambridge Biomedical Campus, Cambridge, UK <sup>51</sup>Department of Public Health and Caring Sciences, Uppsala University, Uppsala, Sweden <sup>52</sup>Section of Biology and Genetics, Department of Neurosciences, Biomedicine and Movement Sciences, University of Verona, Verona, Italy <sup>53</sup>Human Genetics, Wellcome Trust Sanger Institute, Hinxton, UK <sup>54</sup>The National Institute for Health Research Blood and Transplant Research Unit in Donor Health and Genomics, University of Cambridge, Cambridge, UK <sup>55</sup>University Medical Center Groningen, University of Groningen, Department of Cardiology, The Netherlands <sup>56</sup>Center for Human Genetic Research, Massachusetts General Hospital, Boston, Massachusetts, USA <sup>57</sup>Department of Hygiene and Epidemiology, University of Ioannina Medical School, Ioannina, Greece <sup>58</sup>Farr Institute of Health Informatics Research, Institute of Health Informatics, University College London, London, UK <sup>59</sup>Department of Clinical Sciences, Genetic and Molecular Epidemiology Unit, Lund University, Malmö, Sweden <sup>60</sup>Department of Genetics, Physical Anthropology and Animal Physiology, Faculty of Science and Technology, University of the Basque Country (UPV/EHU), Bilbao, Spain <sup>61</sup>University of Glasgow, Glasgow, UK <sup>62</sup>Department of Gerontology and Geriatrics, Leiden University Medical Center, Leiden, The Netherlands <sup>63</sup>Mr. De Craen suddenly passed away January 2016 <sup>64</sup>Department of Cardiology, Leiden University Medical Center, Leiden, The Netherlands <sup>65</sup>Centre for Cognitive Ageing and Cognitive Epidemiology, University of Edinburgh, Edinburgh, UK <sup>66</sup>Centre for Genomic and Experimental Medicine, Medical Research Council Institute of Genetics and Molecular Medicine, University of Edinburgh, Edinburgh, UK <sup>67</sup>Department of Psychology, University of Edinburgh, Edinburgh, UK <sup>68</sup>Queensland Brain Institute, The University of Queensland, Brisbane, Queensland, Australia <sup>69</sup>Department of Twin Research and Genetic Epidemiology, King's College London, UK

<sup>70</sup>Department of Nutrition and Dietetics, School of Health Science and Education, Harokopio University, Athens, Greece <sup>71</sup>Department of Biobank Research, Umeå University, Umeå, Sweden <sup>72</sup>Kuopio Research Institute of Exercise Medicine, Kuopio, Finland <sup>73</sup>Section of Cardiology, Department of Medicine, Boston University Schools of Medicine and Public Health, Boston, Massachusetts, USA <sup>74</sup>Sections of Preventive Medicine and Epidemiology, Department of Medicine, Boston University Schools of Medicine and Public Health, Boston, Massachusetts, USA <sup>75</sup>Department of Epidemiology, Erasmus MC, University Medical Center Rotterdam, Rotterdam, The Netherlands <sup>76</sup>A full list of members and affiliations appears in the Supplementary Note <sup>77</sup>Development Management and Planning, Pfizer Worldwide Research and Development <sup>78</sup>Pfizer Worldwide Research and Development, Stockholm, Sweden <sup>79</sup>Merck Research Laboratories, Genetics and Pharmacogenomics, Boston, Massachusetts, USA <sup>80</sup>Merck Research Laboratories, Cardiometabolic Disease, Kenilworth, New Jersey, USA <sup>81</sup>CHDI Management/CHDI Foundation, Princeton, New Jersey, USA <sup>82</sup>Section of Computational Biomedicine, Department of Medicine, Boston University School of Medicine, Boston, Massachusetts, USA <sup>83</sup>Institute of Epidemiology and Biobank Popgen, Kiel University, Kiel, Germany <sup>84</sup>Neurology Unit, University of Cambridge, Cambridge Biomedical Campus, Cambridge, UK <sup>85</sup>University of North Carolina at Chapel Hill, Department of Epidemiology, Chapel Hill, North Carolina, USA <sup>86</sup>Department of Public Health and Clinical Nutrition, University of Eastern Finland, Finland <sup>87</sup>Research Unit, Kuopio University Hospital, Kuopio, Finland <sup>88</sup>Institute of Biomedicine/Physiology, University of Eastern Finland, Kuopio Campus, Finland <sup>89</sup>Department of Clinical Physiology and Nuclear Medicine, Kuopio University Hospital, Kuopio, Finland <sup>90</sup>Centre for Global Health Research, Usher Institute for Population Health Sciences and Informatics, University of Edinburgh, Edinburgh, UK <sup>91</sup>Faculty of Medicine, University of Split, Croatia <sup>92</sup>Department of Public Health and Clinical Medicine, Umeå University, Umeå, Sweden <sup>93</sup>Department of Nutrition, Harvard School of Public Health, Boston, Massachusetts, USA <sup>94</sup>Alzheimer Scotland Research Centre, University of Edinburgh, Edinburgh, UK <sup>95</sup>Institute of Cardiovascular and Medical Sciences, College of Medical, Veterinary and Life Sciences, University of Glasgow, Glasgow, UK <sup>96</sup>The Interuniversity Cardiology Institute of the Netherlands, Utrecht, The Netherlands <sup>97</sup>Division of Endocrinology, Boston Children's Hospital, Boston, Massachusetts, USA <sup>98</sup>Institute of Molecular and Cell Biology, Tartu, Estonia <sup>99</sup>Department of Cardiology, University of Groningen, University Medical Center Groningen, Groningen, The Netherlands <sup>100</sup>Department of Genetics, University of Groningen, University Medical Center Groningen, Groningen, The Netherlands <sup>101</sup>Durrer Center for Cardiogenetic Research, ICIN-Netherlands Heart Institute, Utrecht, The Netherlands <sup>102</sup>Division of Nephrology, Department of Internal Medicine and Medical Specialties, Columbus - Gemelli University Hospital, Catholic University, Rome, Italy <sup>103</sup>Department of Medical Sciences, Molecular Epidemiology and Science for Life Laboratory, Uppsala University, Uppsala, Sweden <sup>104</sup>Department of Medicine, Division of Cardiovascular Medicine, Stanford University School of Medicine, Stanford, California, USA

<sup>105</sup>Department of Medical Genetics, Center for Molecular Medicine, University Medical Center Utrecht, Utrecht, The Netherlands <sup>106</sup>Department of Epidemiology, Julius Center for Health Sciences and Primary Care, University Medical Center Utrecht, Utrecht, The Netherlands <sup>107</sup>Department of Public Health and Primary Care, Leiden University Medical Center, Leiden, The Netherlands <sup>108</sup>Department of Clinical Biochemistry, Lillebaelt Hospital, Vejle, Denmark <sup>109</sup>Institute of Regional Health Research, University of Southern Denmark, Odense, Denmark <sup>110</sup>Medical Department, Lillebaelt Hospital, Vejle, Denmark <sup>111</sup>Department of Clinical Immunology and Biochemistry, Lillebaelt Hospital, Vejle, Denmark <sup>112</sup>South Ostrobothnia Central Hospital, Seinäjoki, Finland <sup>113</sup>Tampere University Hospital, Tampere, Finland <sup>114</sup>Imperial College Healthcare NHS Trust, London, UK <sup>115</sup>National Heart and Lung Institute, Imperial College London, London, UK <sup>116</sup>Institute of Reproductive and Developmental Biology, Imperial College London, London, UK <sup>117</sup>Department of Epidemiology and Biostatistics, Medical Research Council Public Health England Centre for Environment and Health, School of Public Health, Faculty of Medicine, Imperial College London, St. Mary's Campus, London, UK <sup>118</sup>Centre for Life Course Epidemiology, Faculty of Medicine, University of Oulu, Oulu, Finland <sup>119</sup>Biocenter Oulu, University of Oulu, Oulu, Finland <sup>120</sup>Unit of Primary Care, Oulu University Hospital, Oulu, Finland <sup>121</sup>Research Centre for Prevention and Health, Capital Region of Denmark, Copenhagen, Denmark <sup>122</sup>Department of Clinical Experimental Research, Glostrup University Hospital, Glostrup, Denmark <sup>123</sup>Department of Clinical Medicine, Faculty of Health and Medical Sciences, University of Copenhagen, Copenhagen, Denmark <sup>124</sup>Dasman Diabetes Institute, Dasman, Kuwait <sup>125</sup>Centre for Vascular Prevention, Danube-University Krems, Krems, Austria <sup>126</sup>King Abdulaziz University, Jeddah, Saudi Arabia <sup>127</sup>School of Molecular, Genetic and Population Health Sciences, University of Edinburgh, Medical School, Teviot Place, Edinburgh, UK <sup>128</sup>HUNT Research Centre, Department of Public Health and General Practice, Norwegian University of Science and Technology, Levanger, Norway <sup>129</sup>St. Olav Hospital, Trondheim University Hospital, Trondheim, Norway <sup>130</sup>Department of Medicine, Levanger Hospital, Nord-Trøndelag Health Trust, Levanger, Norway <sup>131</sup>Department of Computational Medicine and Bioinformatics, University of Michigan, Ann Arbor, Michigan, USA <sup>132</sup>Department of Human Genetics, University of Michigan, Ann Arbor, Michigan, USA <sup>133</sup>Folkhälsan Research Centre, Helsinki, Finland; Department of Endocrinology, Helsinki University Central Hospital, Helsinki, Finland <sup>134</sup>Institute for Molecular Medicine Finland University of Helsinki, Helsinki, Finland <sup>135</sup>Department of Clinical Sciences, Diabetes and Endocrinology, Lund University Diabetes Centre, Malmö, Sweden <sup>136</sup>Department of Primary Health Care, Vaasa Central Hospital, Vaasa, Finland <sup>137</sup>Diabetes Center, Vaasa Health Care Center, Vaasa, Finland <sup>138</sup>Psychiatric and Neurodevelopmental Genetics Unit, Department of Psychiatry, Massachusetts General Hospital, Boston, Massachusetts, USA <sup>139</sup>Department of Public Health, University of Helsinki, Finland <sup>140</sup>ICDDR, B; Mohakhali, Dhaka, Bangladesh <sup>141</sup>National Institute of Cardiovascular Diseases, Sher-e-Bangla Nagar, Dhaka, Bangladesh <sup>142</sup>International Centre for Circulatory Health, Imperial College

London, UK <sup>143</sup>Molecular and Cellular Therapeutics, Royal College of Surgeons in Ireland, Dublin, Ireland <sup>144</sup>University of Lille, UMR1167, Risk Factors and Molecular Determinants of aging-related diseases, Lille, France <sup>145</sup>Inserm, Lille, France <sup>146</sup>Centre Hospitalier Universitaire Lille, Public Health, Lille, France <sup>147</sup>Institut Pasteur de Lille, Lille, France <sup>148</sup>Department of Epidemiology and Public Health, EA 3430, University of Strasbourg, Strasbourg, France <sup>149</sup>Department of General and Interventional Cardiology, University Heart Center Hamburg, Germany <sup>150</sup>University Medical Center Hamburg-Eppendorf, Hamburg, Germany <sup>151</sup>Department of Epidemiology, UMR 1027- INSERM, Toulouse University-CHU Toulouse, Toulouse, France <sup>152</sup>Director, UKCRC Centre of Excellence for Public Health, Queens University, Belfast, Northern Ireland <sup>153</sup>Institute of Genetic Epidemiology, Helmholtz Zentrum München - German Research Center for Environmental Health, Neuherberg, Germany <sup>154</sup>Department of Medicine I, University Hospital Grosshadern, Ludwig-Maximilians-Universität, Munich, Germany <sup>155</sup>DZHK (German Centre for Cardiovascular Research), partner site Munich Heart Alliance, Munich, Germany <sup>156</sup>Research Center in Epidemiology and Preventive Medicine, Department of Clinical and Experimental Medicine, University of Insubria, Varese, Italy <sup>157</sup>Princess Al-Jawhara Al-Brahim Centre of Excellence in Research of Hereditary Disorders (PACER-HD), King Abdulaziz University, Jeddah, Saudi Arabia <sup>158</sup>Cardiovascular Research Center, Massachusetts General Hospital, Boston, Massachusetts, USA <sup>159</sup>Department of Medicine, Harvard Medical School, Boston, Massachusetts, USA <sup>160</sup>A Collaboration between the University Medical Schools and NHS, Aberdeen, Dundee, Edinburgh and Glasgow, UK <sup>161</sup>Department of Genetics, University of North Carolina, Chapel Hill, North Carolina, USA <sup>162</sup>Institute of Cardiovascular Sciences, University of Manchester, Manchester, UK <sup>163</sup>Faculty of Population Health Sciences, Institute of Cardiovascular Science, University College London, London, UK <sup>164</sup>The Big Data Institute at the Li Ka Shing Centre for Health Information and Discovery, University of Oxford, Oxford, UK <sup>165</sup>Wellcome Trust Sanger Institute, Hinxton, UK <sup>166</sup>The National Institute for Health Research Blood and Transplant Research

## Acknowledgements

(See Supplementary information)

## References

1. Lim SS, et al. A comparative risk assessment of burden of disease and injury attributable to 67 risk factors and risk factor clusters in 21 regions, 1990-2010: a systematic analysis for the Global Burden of Disease Study 2010. *Lancet*. 2012; 380:2224–60. [PubMed: 23245609]
2. Rapsomaniki E, et al. Blood pressure and incidence of twelve cardiovascular diseases: lifetime risks, healthy life-years lost, and age-specific associations in 1.25 million people. *Lancet*. 2014; 383:1899–911. [PubMed: 24881994]
3. Munroe PB, Barnes MR, Caulfield MJ. Advances in blood pressure genomics. *Circ Res*. 2013; 112:1365–79. [PubMed: 23661711]

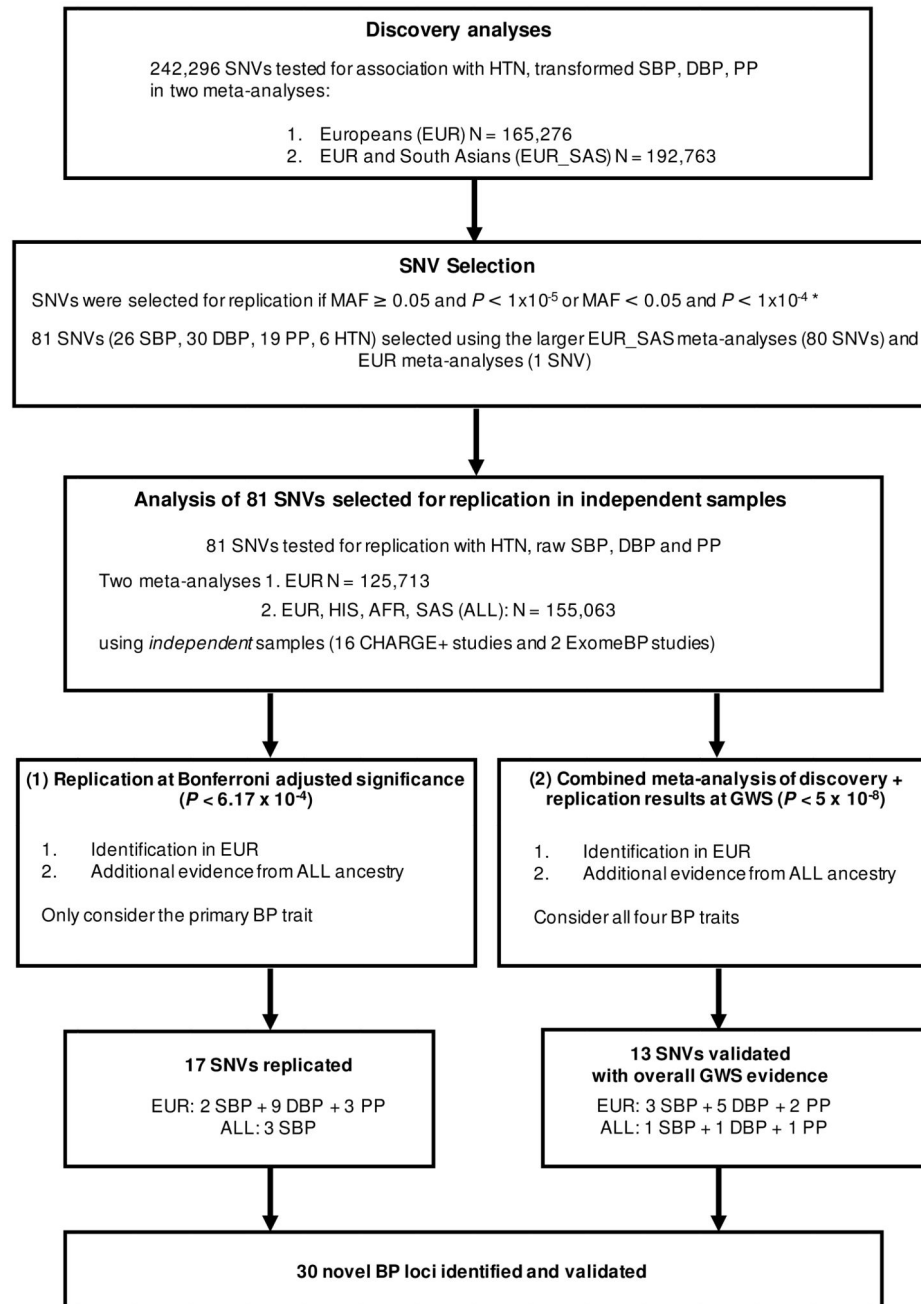
4. Ehret GB, et al. Genetic variants in novel pathways influence blood pressure and cardiovascular disease risk. *Nature*. 2011; 478:103–109. [PubMed: 21909115]
5. Wain LV, et al. Genome-wide association study identifies six new loci influencing pulse pressure and mean arterial pressure. *Nature Genetics*. 2011
6. Johnson T, et al. Blood pressure loci identified with a gene-centric array. *Am J Hum Genet*. 2011; 89:688–700. [PubMed: 22100073]
7. Tomaszewski M, et al. Genetic architecture of ambulatory blood pressure in the general population: insights from cardiovascular gene-centric array. *Hypertension*. 2010; 56:1069–76. [PubMed: 21060006]
8. Tragante V, et al. Gene-centric meta-analysis in 87,736 individuals of European ancestry identifies multiple blood-pressure-related loci. *Am J Hum Genet*. 2014; 94:349–60. [PubMed: 24560520]
9. Ganesh SK, et al. Loci influencing blood pressure identified using a cardiovascular gene-centric array. *Hum Mol Genet*. 2013
10. Simino J, et al. Gene-age interactions in blood pressure regulation: a large-scale investigation with the CHARGE, Global BPgen, and ICBP Consortia. *Am J Hum Genet*. 2014; 95:24–38. [PubMed: 24954895]
11. Zhu X, et al. Meta-analysis of Correlated Traits via Summary Statistics from GWASs with an Application in Hypertension. *Am J Hum Genet*. 2015; 96:21–36. [PubMed: 25500260]
12. Salfati E, Morrison AC, Boerwinkle E, Chakravarti A. Direct Estimates of the Genomic Contributions to Blood Pressure Heritability within a Population-Based Cohort (ARIC). *PLoS One*. 2015; 10:e0133031. [PubMed: 26162070]
13. Schork NJ, Murray SS, Frazer KA, Topol EJ. Common vs. rare allele hypotheses for complex diseases. *Curr Opin Genet Dev*. 2009; 19:212–9. [PubMed: 19481926]
14. Nejentsev S, Walker N, Riches D, Egholm M, Todd JA. Rare variants of IFIH1, a gene implicated in antiviral responses, protect against type 1 diabetes. *Science*. 2009; 324:387–9. [PubMed: 19264985]
15. Liu DJ, et al. Meta-analysis of gene-level tests for rare variant association. *Nat Genet*. 2014; 46:200–4. [PubMed: 24336170]
16. Wu MC, et al. Rare-variant association testing for sequencing data with the sequence kernel association test. *Am J Hum Genet*. 2011; 89:82–93. [PubMed: 21737059]
17. Li B, Leal SM. Methods for detecting associations with rare variants for common diseases: application to analysis of sequence data. *Am J Hum Genet*. 2008; 83:311–21. [PubMed: 18691683]
18. Vissers LE, et al. Heterozygous germline mutations in A2ML1 are associated with a disorder clinically related to Noonan syndrome. *Eur J Hum Genet*. 2015; 23:317–24. [PubMed: 24939586]
19. International Consortium for Blood Pressure Genome-Wide Association, S. Genetic variants in novel pathways influence blood pressure and cardiovascular disease risk. *Nature*. 2011; 478:103–9. [PubMed: 21909115]
20. Coronary Artery Disease Genetics Consortium C4D. A genome-wide association study in Europeans and South Asians identifies five new loci for coronary artery disease. *Nat Genet*. 2011; 43:339–44. [PubMed: 21378988]
21. Schunkert H, et al. Large-scale association analysis identifies 13 new susceptibility loci for coronary artery disease. *Nat Genet*. 2011; 43:333–8. [PubMed: 21378990]
22. Arking DE, et al. Genetic association study of QT interval highlights role for calcium signaling pathways in myocardial repolarization. *Nat Genet*. 2014; 46:826–36. [PubMed: 24952745]
23. den Hoed M, et al. Identification of heart rate-associated loci and their effects on cardiac conduction and rhythm disorders. *Nat Genet*. 2013; 45:621–31. [PubMed: 23583979]
24. Global Lipids Genetics Consortium, et al. Discovery and refinement of loci associated with lipid levels. *Nat Genet*. 2013; 45:1274–83. [PubMed: 24097068]
25. Traylor M, et al. Genetic risk factors for ischaemic stroke and its subtypes (the METASTROKE collaboration): a meta-analysis of genome-wide association studies. *Lancet Neurol*. 2012; 11:951–62. [PubMed: 23041239]



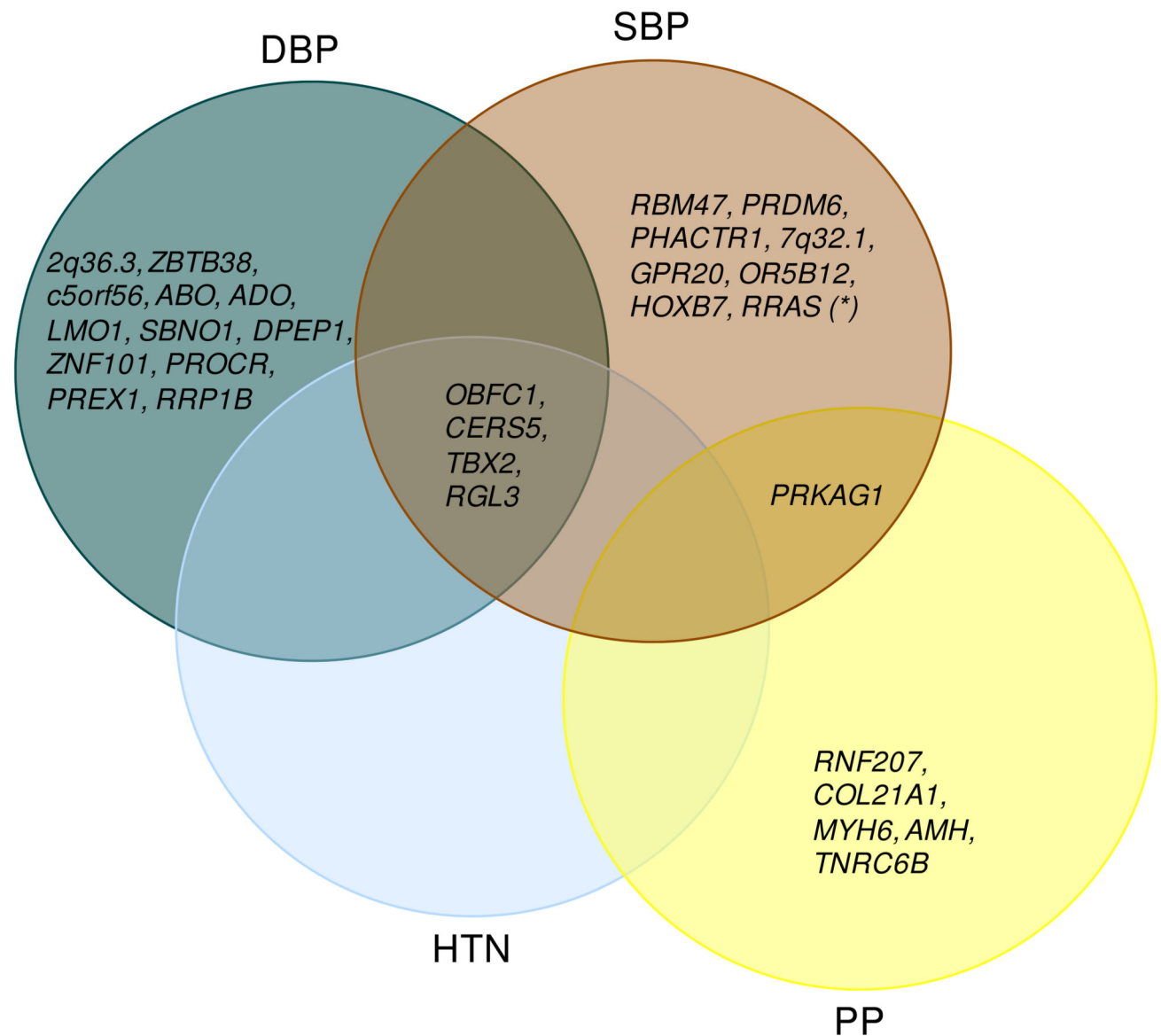
26. Smith NL, et al. Association of genome-wide variation with the risk of incident heart failure in adults of European and African ancestry: a prospective meta-analysis from the cohorts for heart and aging research in genomic epidemiology (CHARGE) consortium. *Circ Cardiovasc Genet*. 2010; 3:256–66. [PubMed: 20445134]
27. Vasan RS, et al. Genetic variants associated with cardiac structure and function: a meta-analysis and replication of genome-wide association data. *JAMA*. 2009; 302:168–78. [PubMed: 19584346]
28. Ehret GB, et al. Genetic variants in novel pathways influence blood pressure and cardiovascular disease risk. *Nature*. 2011; 478:103–109. [PubMed: 21909115]
29. Russo SB, et al. Ceramide synthase 5 mediates lipid-induced autophagy and hypertrophy in cardiomyocytes. *J Clin Invest*. 2012; 122:3919–30. [PubMed: 23023704]
30. Oudot-Mellakh T, et al. Genome wide association study for plasma levels of natural anticoagulant inhibitors and protein C anticoagulant pathway: the MARTHA project. *Br J Haematol*. 2012; 157:230–9. [PubMed: 22443383]
31. Smith NL, et al. Novel associations of multiple genetic loci with plasma levels of factor VII, factor VIII, and von Willebrand factor: The CHARGE (Cohorts for Heart and Aging Research in Genome Epidemiology) Consortium. *Circulation*. 2010; 121:1382–92. [PubMed: 20231535]
32. Bleil ME, Gregorich SE, McConnell D, Rosen MP, Cedars MI. Does accelerated reproductive aging underlie premenopausal risk for cardiovascular disease? *Menopause*. 2013; 20:1139–46. [PubMed: 23591257]
33. Guan R, et al. rbm47, a novel RNA binding protein, regulates zebrafish head development. *Dev Dyn*. 2013; 242:1395–404. [PubMed: 24038582]
34. Wozniak MA, Kwong L, Chodniewicz D, Klemke RL, Keely PJ. R-Ras controls membrane protrusion and cell migration through the spatial regulation of Rac and Rho. *Mol Biol Cell*. 2005; 16:84–96. [PubMed: 15525681]
35. Tuckwell D. Identification and analysis of collagen alpha 1(XXI), a novel member of the FACIT collagen family. *Matrix Biol*. 2002; 21:63–6. [PubMed: 11827793]
36. Huang J, et al. Improved imputation of low-frequency and rare variants using the UK10K haplotype reference panel. *Nat Commun*. 2015; 6:8111. [PubMed: 26368830]
37. Segre AV, et al. Common inherited variation in mitochondrial genes is not enriched for associations with type 2 diabetes or related glycemic traits. *PLoS Genet*. 2010; 6
38. Huan T, et al. Integrative network analysis reveals molecular mechanisms of blood pressure regulation. *Mol Syst Biol*. 2015; 11:799.
39. Locke AE, et al. Genetic studies of body mass index yield new insights for obesity biology. *Nature*. 2015; 518:197–206. [PubMed: 25673413]
40. Wood AR, et al. Defining the role of common variation in the genomic and biological architecture of adult human height. *Nat Genet*. 2014; 46:1173–86. [PubMed: 25282103]
41. Park HW, et al. Serine-threonine kinase with-no-lysine 4 (WNK4) controls blood pressure via transient receptor potential canonical 3 (TRPC3) in the vasculature. *Proc Natl Acad Sci U S A*. 2011; 108:10750–5. [PubMed: 21670282]
42. Sudlow C, et al. UK biobank: an open access resource for identifying the causes of a wide range of complex diseases of middle and old age. *PLoS Med*. 2015; 12:e1001779. [PubMed: 25826379]
43. Te Riet L, van Esch JH, Roks AJ, van den Meiracker AH, Danser AH. Hypertension: Renin-Angiotensin-aldosterone system alterations. *Circ Res*. 2015; 116:960–75. [PubMed: 25767283]
44. Gao J, et al. A new strategy for treating hypertension by blocking the activity of the brain renin-angiotensin system with aminopeptidase A inhibitors. *Clin Sci (Lond)*. 2014; 127:135–48. [PubMed: 24697296]
45. Tobin MD, Sheehan NA, Scurrah KJ, Burton PR. Adjusting for treatment effects in studies of quantitative traits: antihypertensive therapy and systolic blood pressure. *Stat Med*. 2005; 24:2911–35. [PubMed: 16152135]
46. Mahajan A, et al. Identification and Functional Characterization of G6PC2 Coding Variants Influencing Glycemic Traits Define an Effector Transcript at the G6PC2-ABCB11 Locus. *PLoS Genet*. 2015; 11:e1004876. [PubMed: 25625282]
47. Willer CJ, Li Y, Abecasis GR. METAL: fast and efficient meta-analysis of genomewide association scans. *Bioinformatics*. 2010; 26:2190–1. [PubMed: 20616382]



48. Burgess S, Thompson SG. Multivariable Mendelian randomization: the use of pleiotropic genetic variants to estimate causal effects. *Am J Epidemiol.* 2015; 181:251–60. [PubMed: 25632051]
49. Thompson SG, Sharp SJ. Explaining heterogeneity in meta-analysis: a comparison of methods. *Stat Med.* 1999; 18:2693–708. [PubMed: 10521860]
50. Nica AC, et al. The architecture of gene regulatory variation across multiple human tissues: the MuTHER study. *PLoS Genet.* 2011; 7:e1002003. [PubMed: 21304890]
51. GTEx, Consortium. Human genomics. The Genotype-Tissue Expression (GTEx) pilot analysis: multitissue gene regulation in humans. *Science.* 2015; 348:648–60. [PubMed: 25954001]
52. Giambartolomei C, et al. Bayesian test for colocalisation between pairs of genetic association studies using summary statistics. *PLoS Genet.* 2014; 10:e1004383. [PubMed: 24830394]

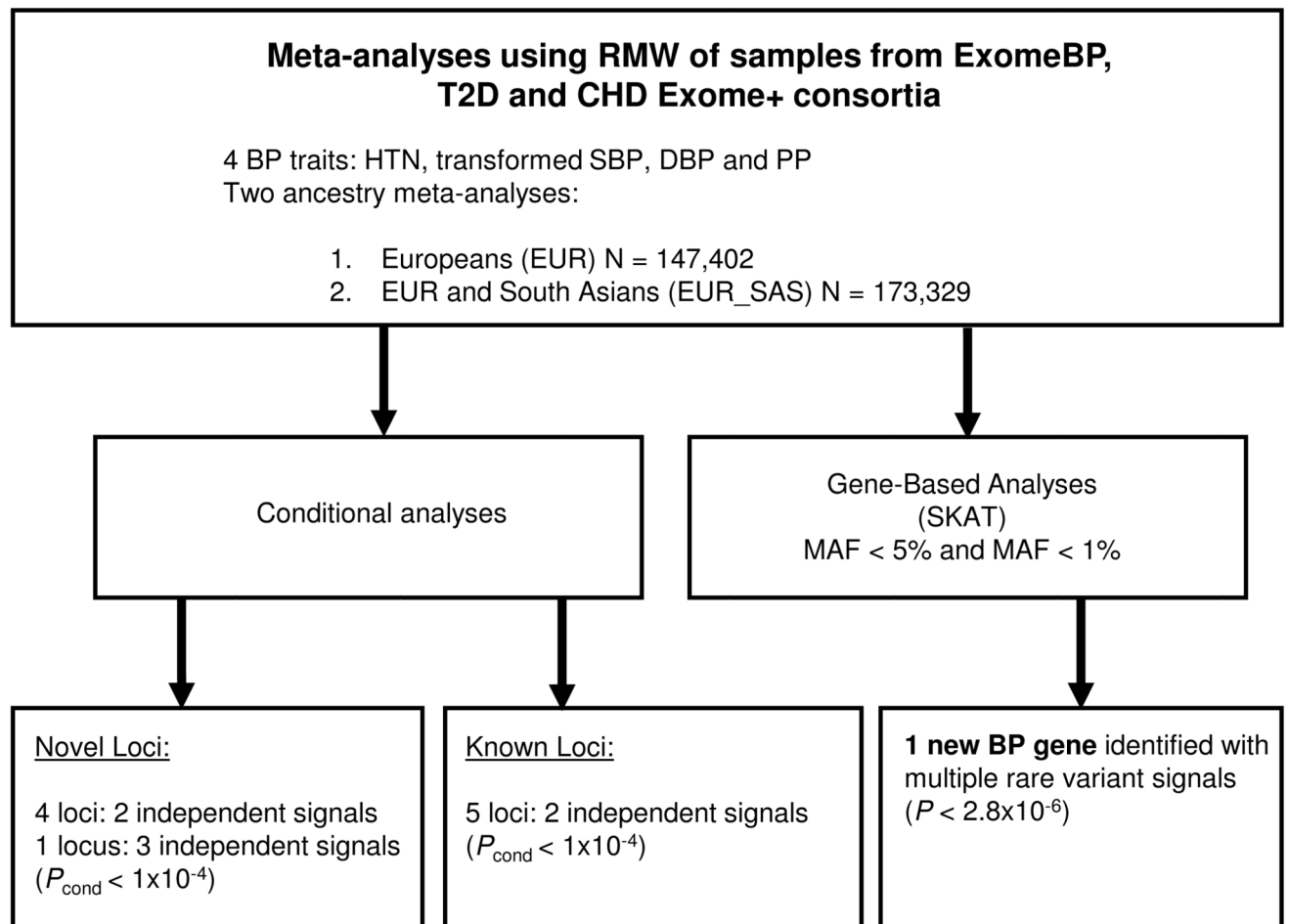


**Figure 1. Study design and work flow diagram of single variant discovery analyses.**  
EUR=European, SAS=South Asian, HIS=Hispanic, AA=African American,  
HTN=hypertension, BP=blood pressure, SBP=systolic blood pressure, DBP= diastolic blood  
pressure, PP=pulse pressure, N=sample size, MAF=minor allele frequency,  $P$ = $P$ -value  
significance threshold, SNV=single-nucleotide variant, GWS=genome-wide significance  
\*Further details of the selection criteria are provided in the methods.

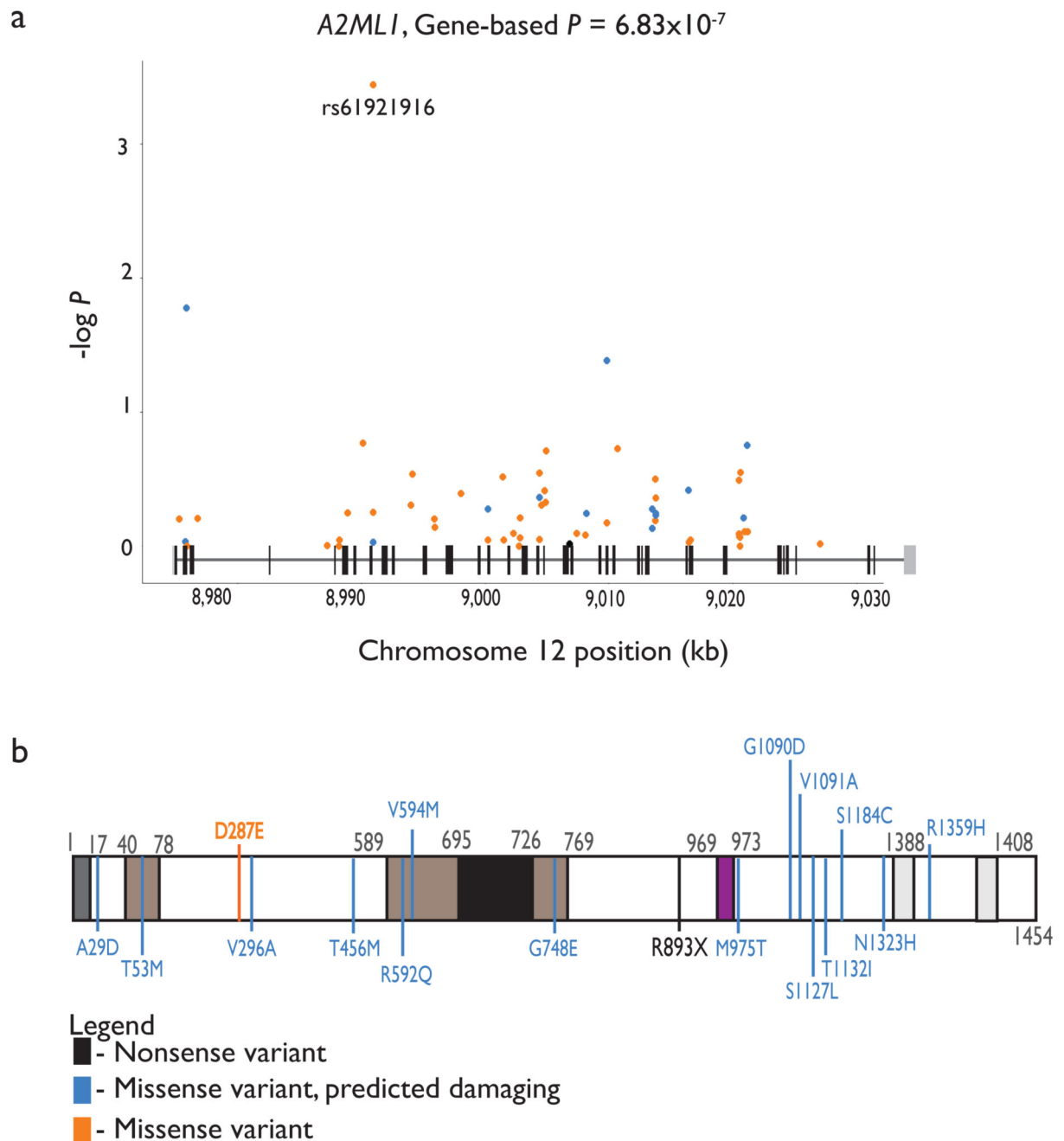


**Figure 2. Overlap of the 30 novel loci associations across SBP, DBP, PP and HTN.**

The Venn diagram shows which of the 30 newly identified BP loci are associated with multiple BP traits. Only SNV-BP trait associations that were genome-wide significant ( $P < 5 \times 10^{-8}$ ) in the combined discovery and replication meta-analyses are listed for any given BP trait, within the corresponding ancestry dataset that the given locus was validated for (see Tables 1 and 2). The association of *RRAS* variant with SBP was replicated in the independent samples, but did not achieve GWS in the combined discovery and replication meta-analysis and is therefore only included for SBP. HTN=hypertension, SBP=systolic blood pressure, DBP= diastolic blood pressure, PP=pulse pressure.



**Figure 3. Study design for conditional analyses and rare variant gene-based discovery analyses.** RMW=RareMetalWorker, EUR=European, SAS = South Asian, HTN=hypertension, BP=blood pressure, SBP=systolic blood pressure, DBP= diastolic blood pressure, PP=pulse pressure. N=sample size, MAF=minor allele frequency,  $P$ = $P$ -value significance threshold,  $P_{\text{cond}}$ =conditional  $P$ -value significance threshold



**Figure 4. Locus plot for *A2ML1* and secondary amino acid structure of the gene product.**

(a) Locus plot for *A2ML1* associated with HTN identified through gene based tests. The variants' positions along the gene (x axis, based on human genome build 37) and the  $-\log_{10}(P\text{-value of association})$  (y axis) are indicated. The variants are colour coded: nonsense (black), missense, predicted damaging (blue), and missense (orange). The schematic above the x-axis represents the intron / exon (black vertical bars) structure, the untranslated regions are shown as grey vertical bars.

(b) The white box denotes the full-length amino acid sequence for each of the two gene products. Black numbers denote amino acid residue positions of note. Coloured boxes depict putative functional domains (see below). Coloured vertical lines indicate the amino acid substitutions corresponding to the variants depicted in the locus plots above using the same colour coding. Bold, italic indicates the SNV association with smallest  $P$ -value. Dark grey – signal peptide sequence. Brown – regions of intramolecular disulfide bonds. For simplicity only those regions coinciding with variants described were indicated. Black – bait region described to interact with proteases. Purple – thiol ester sequence region aiding in interaction with proteases. Light grey – alpha helical regions thought to mediate A2ML1 interaction with LRP1, facilitating receptor-mediated endocytosis.



**Table 1**  
**Novel blood pressure trait associated loci. Variants with formal replication**

Variant information				Discovery		Replication			Combined		
Locus	rsID	Chr:Pos (EA:EAF)	Trait	$P_t$	$P_U$	N	$\beta$	$P$	N	$\beta$	$P$
<b>EUR</b>											
<i>RNF207</i>	rs709209	1:6.28 (A:0.655)	PP	$4.57 \times 10^{-6}$	$1.60 \times 10^{-6}$	122,780	0.17	$5.83 \times 10^{-4}$	284,683	0.20	$9.62 \times 10^{-9}$
<i>C5orf56</i>	rs12521868	5:131.78 (T:0.373)	DBP	$1.59 \times 10^{-6}$	$3.03 \times 10^{-7}$	122,795	-0.18	$2.29 \times 10^{-5}$	282,023	-0.19	$6.12 \times 10^{-11}$
<i>PHACTR1</i>	rs9349379	6:12.90 (A:0.566)	SBP	$2.11 \times 10^{-8}$	$1.78 \times 10^{-7}$	122,809	0.24	$4.06 \times 10^{-4}$	284,673	0.29	$8.84 \times 10^{-10}$
<i>COL21A1</i>	rs200999181 <sup>†</sup>	6:55.94 (A:0.002)	PP	$3.08 \times 10^{-8}$	$2.46 \times 10^{-7}$	121,487	2.70	$1.90 \times 10^{-4}$	242,486	3.25	$6.27 \times 10^{-10}$
<i>ABO</i>	rs687621	9:136.14 (A:0.615)	DBP	$8.80 \times 10^{-8}$	$2.55 \times 10^{-7}$	122,798	0.16	$1.96 \times 10^{-4}$	276,014	0.19	$5.45 \times 10^{-10}$
<i>ADO</i>	rs10995311	10:64.56 (C:0.567)	DBP	$1.86 \times 10^{-6}$	$1.14 \times 10^{-6}$	122,798	0.23	$8.47 \times 10^{-8}$	266,456	0.21	$1.12 \times 10^{-12}$
<i>LMO1</i>	rs110419	11:8.25 (A:0.48)	DBP	$9.41 \times 10^{-6}$	$2.22 \times 10^{-5}$	122,798	0.16	$1.81 \times 10^{-4}$	279,935	0.16	$3.04 \times 10^{-8}$
<i>OR5B12</i>	rs11229457	11:58.21 (T:0.236)	SBP	$1.58 \times 10^{-6}$	$4.62 \times 10^{-5}$	122,809	-0.32	$7.53 \times 10^{-5}$	284,680	-0.31	$2.70 \times 10^{-8}$
<i>CERS5</i>	rs7302981	12:50.54 (A:0.361)	DBP	$1.35 \times 10^{-13}$	$4.60 \times 10^{-11}$	122,798	0.24	$2.64 \times 10^{-8}$	284,718	0.25	$1.38 \times 10^{-17}$
<i>MYH6</i>	rs452036	14:23.87 (A:0.327)	PP	$4.59 \times 10^{-11}$	$2.80 \times 10^{-13}$	122,780	-0.21	$1.81 \times 10^{-5}$	284,672	-0.28	$2.96 \times 10^{-16}$
<i>DPEP1</i>	rs1126464	16:89.70 (C:0.256)	DBP	$1.19 \times 10^{-9}$	$4.35 \times 10^{-11}$	118,677	0.24	$1.68 \times 10^{-6}$	261,564	0.28	$1.02 \times 10^{-15}$
<i>TBX2</i>	rs8068318 <sup>†</sup>	17:59.48 (T:0.698)	DBP	$7.46 \times 10^{-13}$	$5.71 \times 10^{-10}$	122,798	0.26	$3.23 \times 10^{-8}$	281,978	0.26	$1.95 \times 10^{-16}$
<i>RGL3</i>	rs167479	19:11.53 (T:0.486)	DBP	$2.22 \times 10^{-23}$	$1.97 \times 10^{-22}$	122,797	-0.29	$3.01 \times 10^{-11}$	283,332	-0.33	$1.99 \times 10^{-31}$
<i>PREX1</i>	rs6095241	20:47.31 (A:0.452)	DBP	$5.65 \times 10^{-6}$	$2.29 \times 10^{-5}$	122,798	-0.18	$2.56 \times 10^{-5}$	281,322	-0.17	$4.75 \times 10^{-9}$
<b>ALL ancestry</b>											
<i>RBM47</i>	rs35529250 <sup>†</sup>	4:40.43 (T:0.01)	SBP	$6.56 \times 10^{-7}$	$6.15 \times 10^{-6}$	148,878	-1.43	$5.02 \times 10^{-4}$	306,352	-1.55	$2.42 \times 10^{-8}$
<i>OBFC1</i>	rs4387287	10:105.68 (A:0.157)	SBP	$2.23 \times 10^{-8}$	$1.32 \times 10^{-7}$	147,791	0.28	$3.37 \times 10^{-4}$	320,494	0.36	$9.12 \times 10^{-10}$
<i>RRAS</i>	rs61760904 <sup>†</sup>	19:50.14 (T:0.008)	SBP	$1.96 \times 10^{-6}$	$1.90 \times 10^{-5}$	148,878	1.38	$5.70 \times 10^{-4}$	322,664	1.50	$8.45 \times 10^{-8}$

SNV-BP associations are reported for the newly identified BP loci that replicated at  $P < 6.2 \times 10^{-4}$  (Bonferroni correction for the 81 variants selected for replication for a primary blood pressure trait; Methods). Loci are categorised into EUR and ALL ancestry based on the meta-analysis used to replicate the variants for the primary BP trait shown in columns labelled 'Trait'. In the columns that contains the discovery meta-analyses results,  $P_t$  represents the  $P$ -value for association of the variant with the transformed primary BP trait in the EUR\_SAS discovery meta-analyses (which was also used to select the variant for replication) and  $P_U$  represents the  $P$ -value for association with the untransformed primary BP trait in the ancestry in which the variant replicated. N,  $\beta$  and  $P$ , which denote the number of samples, estimated allelic effect and  $P$ -value for association with the primary BP trait, are provided for the untransformed primary BP trait in the replication data and also from the combined (discovery and replication) meta-analyses. NB: ALL ancestry corresponds to all ancestries in the combined (discovery + replication) meta-analyses

Locus – Gene or region containing the SNV, rsID – dbSNP rsID. Chr:Pos (EA:EAF) – Chromosome:NCBI Build 37 position in Mb (effect allele:effect allele frequency), Trait – primary blood pressure trait for which the variant was and also replicated,  $\beta$  – effect estimate, N:sample size, EUR – European.

<sup>†</sup> indicates it is a non-synonymous SNV (nsSNV) or is in linkage disequilibrium with a nsSNV ( $r^2 > 0.8$ ) that is predicted to be damaging

**Table 2**  
**Novel blood pressure trait associated loci. Variants with GWS evidence of association in combined meta-analyses**

Variant information				Discovery		Replication			Combined		
Locus	rsID	Chr:Pos (EA:EAF)	Trait	$P_t$	$P_U$	N	$\beta$	$P$	N	$\beta$	$P$
<b>EUR</b>											
2q36.3	rs2972146	2:227.10 (T:0.652)	DBP <sup>§</sup> (HTN)	1.51 x 10 <sup>-9</sup>	2.47 x 10 <sup>-7</sup>	122,798	0.13	2.20 x 10 <sup>-3</sup>	275,610	0.17	8.40 x 10 <sup>-9</sup>
<i>ZBTB38</i>	rs16851397	3:141.13 (A:0.953)	DBP <sup>§</sup> (SBP)	6.87 x 10 <sup>-6</sup>	3.20 x 10 <sup>-5</sup>	122,798	-0.38	1.20 x 10 <sup>-4</sup>	284,717	-0.38	3.01 x 10 <sup>-8</sup>
<i>PRDM6</i>	rs1008058	5:122.44 (A:0.135)	SBP	5.09 x 10 <sup>-7</sup>	1.01 x 10 <sup>-8</sup>	43,109	0.46	3.61 x 10 <sup>-3</sup>	176,362	0.55	2.99 x 10 <sup>-10</sup>
<i>GPR20</i>	rs34591516	8:142.37 (T:0.055)	SBP <sup>§</sup> (DBP)	1.54 x 10 <sup>-6</sup>	1.01 x 10 <sup>-7</sup>	122,807	0.51	4.20 x 10 <sup>-4</sup>	282,009	0.64	6.10 x 10 <sup>-10</sup>
<i>HOXB7</i>	rs7406910	17:46.69 (T:0.118)	SBP	6.07 x 10 <sup>-10</sup>	2.74 x 10 <sup>-9</sup>	122,809	-0.20	4.89 x 10 <sup>-2</sup>	284,690	-0.46	3.80 x 10 <sup>-8</sup>
<i>AMH</i>	rs10407022 <sup>‡</sup>	19:2.25 (T:0.82)	PP	1.63 x 10 <sup>-7</sup>	1.73 x 10 <sup>-7</sup>	118,656	-0.19	1.62 x 10 <sup>-3</sup>	252,525	-0.26	5.94 x 10 <sup>-9</sup>
<i>ZNF101</i>	rs2304130	19:19.79 (A:0.914)	DBP	1.66 x 10 <sup>-8</sup>	1.92 x 10 <sup>-8</sup>	122,798	-0.17	1.71 x 10 <sup>-2</sup>	284,705	-0.29	1.53 x 10 <sup>-8</sup>
<i>PROCR</i>	rs867186	20:33.76 (A:0.873)	DBP	1.44 x 10 <sup>-6</sup>	4.15 x 10 <sup>-7</sup>	122,798	0.21	2.48 x 10 <sup>-3</sup>	284,722	0.26	1.19 x 10 <sup>-8</sup>
<i>RRP1B</i>	rs9306160	21:45.11 (T:0.374)	DBP <sup>§</sup> (SBP)	1.04 x 10 <sup>-8</sup>	1.90 x 10 <sup>-6</sup>	100,489	-0.16	4.30 x 10 <sup>-4</sup>	249,817	-0.18	6.80 x 10 <sup>-9</sup>
<i>TNRC6B</i>	rs470113	22:40.73 (A:0.804)	PP	1.48 x 10 <sup>-10</sup>	1.31 x 10 <sup>-9</sup>	122,780	-0.14	1.37 x 10 <sup>-2</sup>	284,683	-0.25	1.67 x 10 <sup>-9</sup>
<b>ALL ancestry</b>											
7q32.1	rs4728142	7:128.57 (A:0.433)	SBP	8.10 x 10 <sup>-6</sup>	4.21 x 10 <sup>-6</sup>	150,542	-0.21	8.62 x 10 <sup>-4</sup>	338,338	-0.24	3.45 x 10 <sup>-8</sup>
<i>PRKAG1</i>	rs1126930 <sup>‡</sup>	12:49.40 (C:0.036)	PP	2.12 x 10 <sup>-6</sup>	4.62 x 10 <sup>-7</sup>	151,481	0.36	3.74 x 10 <sup>-3</sup>	314,894	0.50	3.34 x 10 <sup>-8</sup>
<i>SBNO1</i>	rs1060105	12:123.81 (T:0.209)	DBP	6.66 x 10 <sup>-7</sup>	1.09 x 10 <sup>-6</sup>	150,532	-0.15	2.67 x 10 <sup>-3</sup>	336,413	-0.18	3.07 x 10 <sup>-8</sup>

SNV-BP associations are reported for the newly identified BP loci that showed genome-wide significant association ( $P < 5 \times 10^{-8}$ ) in the combined discovery and replication meta-analyses. In the columns that contain results from the discovery meta-analyses,  $P_t$  represents the  $P$ -value for association of the variant with the transformed *primary* BP trait in the EUR\_SAS discovery meta-analyses (used to select the variant for replication) and  $P_U$  represents the  $P$ -value for association with the untransformed BP trait in the ancestry in which the variant was validated. Loci are categorised into EUR and ALL ancestry based on the ancestry in which the variant showed association with a blood pressure trait at  $P < 5 \times 10^{-8}$ . N,  $\beta$  and  $P$ , which denote the number of samples, estimated allelic effect and  $P$ -value for association with the validated BP trait, are provided for the untransformed BP trait in the replication data and also from the combined (discovery and replication) meta-analyses. NB: ALL ancestry corresponds to all ancestries in the combined (discovery + replication) meta-analyses.

Locus – Gene or region containing the SNV, rsID – dbSNP rsID. Chr:Pos (EA:EAF) – Chromosome:NCBI Build 37 position in Mb (effect allele:effect allele frequency), Trait – blood pressure trait for which association is reported, EUR – European.

<sup>§</sup> At four loci (2q36.3, *ZBTB38*, *GPR20* and *RRP1B*) the primary trait used to select the variants for replication is given in parentheses because the variant associations were validated in the combined meta-analysis for the listed secondary trait. For these variants,  $P_t$  denotes the  $P$ -value for association with the primary trait, the other  $P$ -values provided are for the secondary trait.

<sup>‡</sup> indicates it is a non-synonymous SNV (nsSNV) or is linkage disequilibrium with a nsSNV ( $r^2 > 0.8$ ) that is predicted to be damaging

**Table 3**  
**Results of the genetic risk score analyses across CVD traits and risk factors.**

Outcome	Units	N	DBP (per 10mmHg increase)	Effect [95% CI]	P	SBP (per 10mmHg increase)	Effect [95% CI]	P	PP (per 10mmHg increase)	Effect [95% CI]	P
CHD	OR	82,056	1.62 [ 1.28, 2.05]	5.99 x 10 <sup>-5</sup>	1.39 [ 1.22, 1.59]	6.07 x 10 <sup>-7</sup>	1.70 [ 1.34, 2.16]	1.20 x 10 <sup>-5</sup>			
Ischemic stroke	OR	25,799	1.93 [ 1.47, 2.55]	2.81 x 10 <sup>-6</sup>	1.57 [ 1.35, 1.84]	1.16 x 10 <sup>-8</sup>	2.12 [ 1.58, 2.84]	5.35 x 10 <sup>-7</sup>			
Cardioembolic stroke	OR	16,113	1.43 [ 0.86, 2.39]	0.1683	1.33 [ 0.99, 1.80]	0.0584	1.73 [ 1.00, 3.02]	0.0518			
Large vessel stroke	OR	13,903	2.26 [ 1.25, 4.08]	0.0068	1.85 [ 1.32, 2.59]	3.61 x 10 <sup>-4</sup>	3.05 [ 1.64, 5.68]	4.37 x 10 <sup>-4</sup>			
Small vessel stroke	OR	15,617	1.96 [ 1.13, 3.41]	0.0168	1.56 [ 1.13, 2.16]	0.0064	1.98 [ 1.09, 3.61]	0.0248			
Heart failure	OR	13,282	1.48 [ 1.02, 2.17]	0.0409	1.25 [ 1.00, 1.57]	0.0512	1.33 [ 0.88, 2.02]	0.1757			
Left ventricular mass	g	11,273	9.57 [ 3.98, 15.17]	8.02 x 10 <sup>-4</sup>	5.13 [ 1.77, 8.48]	0.0027	5.97 [ -0.38, 12.31]	0.0653			
Left ventricular wall thickness	cm	11,311	0.10 [ 0.06, 0.13]	1.88 x 10 <sup>-8</sup>	0.05 [ 0.03, 0.07]	5.52 x 10 <sup>-6</sup>	0.05 [ 0.01, 0.09]	0.0187			
HDL	mg/dl	80,395	0.25 [ -1.00, 1.51]	0.6930	0.21 [ -0.50, 0.92]	0.5622	0.47 [ -0.79, 1.73]	0.4668			
LDL	mg/dl	77,021	-1.57 [ -5.20, 2.06]	0.3972	0.07 [ -2.03, 2.16]	0.9498	1.87 [ -1.86, 5.59]	0.3255			
Total cholesterol	mg/dl	80,455	-1.34 [ -5.90, 3.22]	0.5639	0.70 [ -1.93, 3.32]	0.6029	3.68 [ -0.97, 8.33]	0.1209			
Triglycerides	mg/dl	77,779	0.02 [ -0.03, 0.08]	0.3859	0.02 [ -0.01, 0.05]	0.2697	0.03 [ -0.03, 0.08]	0.3025			
BMI	INVT	526,508	-0.10 [ -0.18, -0.01]	0.0342	-0.07 [ -0.13, -0.02]	0.0058	-0.12 [ -0.23, -0.02]	0.0165			
WHRadjBMI	INVT	344,369	0.03 [ -0.04, 0.11]	0.4025	0.03 [ -0.02, 0.08]	0.2170	0.06 [ -0.03, 0.15]	0.1885			
Height	INVT	458,927	0.02 [ -0.15, 0.18]	0.8592	-0.04 [ -0.15, 0.06]	0.4170	-0.18 [ -0.37, 0.01]	0.0683			
eGFR	INVT	51,039	-0.02 [ -0.15, 0.11]	0.7810	-0.03 [ -0.10, 0.04]	0.4080	-0.07 [ -0.20, 0.06]	0.2741			

CHD, coronary heart disease; HDL, high density lipoprotein; LDL, low density lipoprotein; eGFR, estimated glomerular filtration rate; DBP, diastolic blood pressure; SBP, systolic blood pressure; PP, pulse pressure; OR, odds ratio; g, grams; INVT, inverse normally transformed (hence no units); N, sample size; P, P-value of association of BP with the trait listed; CI, confidence interval. Results are considered significant if  $P < 0.0038$ , which corresponds to a Bonferroni correction for 13 phenotypes tested.